

Executive Steering Committee
For A.C.E. Policy II
(ESCAP II)
Report 5

March 13, 2002

ESCAP II

E-Sample Erroneous Enumerations

Roxanne Feldpausch

Decennial Statistical
Studies Division

U S C E N S U S B U R E A U

Helping You Make Informed Decisions

CONTENTS

EXECUTIVE SUMMARY	v
1. BACKGROUND	1
1.1 Operational background	1
1.2 Definitions	2
2. METHODS	3
3. LIMITS	3
4. RESULTS	4
4.1 The Different Types of Erroneous Enumerations	4
4.1.1 Race/Hispanic origin	5
4.1.2 Age/Sex	6
4.1.3 Tenure (Owner vs. Non-Owner)	7
4.1.4 Return rate <i>indicator</i>	7
4.1.5 Place size and type of enumeration	8
4.1.6 Region	9
4.1.7 Response method	10
4.1.9 Characteristic Imputation	11
4.1.10 Housing unit's match status	11
4.1.11 Mail Return Date	13
4.1.12 Other variables	15
4.2 What is the difference in erroneous enumeration between 1990 and 2000	16
4.2.1 Changes between 1990 and 2000 that affect erroneous enumeration rates	16
4.2.2 A comparison of erroneous enumeration rates between 1990 and 2000 ..	20
4.2.3 Reclassification of people with unresolved status	20
SUMMARY	21
References	23
Appendix A	
Detailed Tables of Types of Erroneous Enumerations, Standard Errors, Results of Significance Testing and Percent of the E-Sample	25
Appendix B	
Technical Documentation	82

TABLES

Table 1 Types Erroneous Enumerations in Census 2000 (standard errors)	4
Table 2 Percent Erroneous Enumerations by Domain (standard errors)	6
Table 3 Percent Erroneous Enumerations by Age/Sex (standard errors)	7
Table 4 Percent Erroneous Enumerations by Tenure (standard errors)	7
Table 5 Percent Erroneous Enumerations by Return Rate (standard errors)	8
Table 6 Percent Erroneous Enumerations by Metropolitan Statistical Area/Type of Enumeration Area (standard errors)	9
Table 7 Percent Erroneous Enumerations by Region (standard errors)	9
Table 8 Percent Erroneous Enumerations by Response Method (standard errors)	11
Table 9 Percent Erroneous Enumerations by Characteristic Imputation (standard errors)	11
Table 10 Percent Erroneous Enumerations by Housing Unit's Match Status (standard errors)	12
Table 11 Type of Erroneous Enumeration by Initial Address Status (standard errors)	13
Table 12 Percent Erroneous Enumerations Before and After TES	20
Table 13 Comparison Between 1990 and 2000 of the Percent Erroneous Enumerations	20
Table 14: Comparison Between 1990 and 2000 of the Percent Erroneous Enumerations After Re-classifying the Unresolved People	21
Table A-1.A Types of Erroneous Enumerations by Domain	25
Table A-1.B Standard Errors of Types of Erroneous Enumerations by Domain	26
Table A-1.C Significant Differences among Domain	27
Table A-1.D Percent of the E-Sample by Domain	29
Table A-2.A Types of Erroneous Enumerations by Age/Sex	30
Table A-2.B Standard Errors of Types of Erroneous Enumerations by	31
Table A-2.C Significant Differences among Age/Sex	31
Table A-2.D Percent of the E-Sample by Age/Sex	33
Table A-3.A Types of Erroneous Enumerations by Tenure	34
Table A-3.B Standard Errors of Types of Erroneous Enumerations by Tenure	34
Table A-3.C Significant Differences among Tenure	34
Table A-3.D Percent of the E-Sample by Tenure	35
Table A-4.A Types of Erroneous Enumerations by Return Rate	36
Table A-4.B Standard Errors of Types of Erroneous Enumerations by Return Rate	36
Table A-4.C Significant Differences among Return Rate	36
Table A-4.D Percent of the E-Sample by Return Rate	37
Table A-5.A Types of Erroneous Enumerations by MSA/TEA	38
Table A-5.B Standard Errors of Types of Erroneous Enumerations by MSA/TEA	38
Table A-5.C Significant Differences among MSA/TEA	39
Table A-5.D Percent of the E-Sample by MSA/TEA	39
Table A-6.A Types of Erroneous Enumerations by Region	40
Table A-6.B Standard Errors of Types of Erroneous Enumerations by Region	40
Table A-6.C Significant Differences among Regions	41
Table A-6.D Percent of the E-Sample by Region	41
Table A-7.A Types of Erroneous Enumerations by Response Method	42

Table A-7.B Standard Errors of Types of Erroneous Enumerations by Response Method	42
Table A-7.C Significant Differences among Response Methods (Self-reporting, Enumerator returns completed with a household member and enumerator returns complete with a proxy respondent)	43
Table A-7.D Percent of the E-Sample by Response Method	43
Table A-8.A Types of Erroneous Enumerations by Imputation	44
Table A-8.B Standard Errors of Types of Erroneous Enumerations by Imputation	44
Table A-8.C Significant Differences among Imputation	44
Table A-8.D Percent of the E-Sample by Imputation	45
Table A-9.A Types of Erroneous Enumerations by Housing Unit Match Status	46
Table A-9.B Standard Error of Types of Erroneous Enumerations by Housing Unit Match Status	46
Table A-9.C Significant Differences among Housing Unit Match Status	47
Table A-9.D Percent of the E-Sample by Housing Unit Match Status	47
Table A-10.A Type of Erroneous Enumeration by Initial Address Status	48
Table A-10.B Standard Errors of Type of Erroneous Enumeration by Initial Address Status . .	49
Table A-10.C Significant Differences among Initial Address Status	49
Table A-10.D Percent of the E-Sample by initial Address Status	50
Table A-11.A Types of Erroneous Enumerations by Number of Units at the Basic Street Address	51
Table A-11.B Standard Error of Types of Erroneous Enumerations by Number of Units at the Basic Street Address	51
Table A-11.C Significant Differences among number of units at the basic street address	51
Table A-11.D Percent of the E-Sample by Units at Basic Street Address	52
Table A-12.A Types of Erroneous Enumerations by Whole/Partial Match Code	53
Table A-12.B Standard Errors of Types of Erroneous Enumerations by Whole/Partial Match Code	53
Table A-12.C Percent of the E-Sample by Whole/Partial Match Code	54
Table A-13.A Types of Erroneous Enumerations by Form Length	55
Table A-13.B Standard Errors of Types of Erroneous Enumerations by Form Length	55
Table A-13.C Significant Differences among Form Length	55
Table A-13.D Percent of the E-Sample by Form Length	56
Table A-14.A Types of Erroneous Enumerations by Non-Response Follow Up	57
Table A-14.B Standard Errors of Types of Erroneous Enumerations by Non-Response Follow Up	57
Table A-14.C Significant Differences among Non-Response Follow Up	58
Table A-14.D Percent of the E-Sample by Non-Response Follow Up	58
Table A-15.A Type of Erroneous Enumeration by Coverage Edit Follow Up	59
Table A-15.B Standard Errors of Type of Erroneous Enumeration by Coverage Edit Follow Up	59
Table A-15.C Significant Differences among Coverage Edit Follow Up	59
Table A-15.D Percent of the E-Sample by Coverage Edit Follow Up	60
Table A-16.A Types of Erroneous Enumerations by Coverage Improvement Follow Up	61

Table A-16.B Standard Errors of Types of Erroneous Enumerations by Coverage Improvement Follow Up	61
Table A-16.C Significant Differences among Coverage Improvement Follow Up	61
Table A-16.D Percent of the E-Sample by Coverage Improvement Follow Up	62
Table A-17.A Types of Erroneous Enumerations by Capture Method	63
Table A-17.B Standard Errors of Types of Erroneous Enumerations by Capture Method	63
Table A-17.C Significant Differences among Capture Method	63
Table A-17.D Percent of the E-Sample by Capture System	64
Table A-18.A Types of Erroneous Enumerations by Data Capture Center	65
Table A-18.B Standard Errors of Types of Erroneous Enumerations by Data Capture Center ..	65
Table A-18.C Percent of the E-Sample by Data Capture Center	66
Table A-19.A Type of Erroneous Enumerations by Follow-Up	67
Table A-19.B Standard Errors of Type of Erroneous Enumerations by	67
Table A-19.C Percent of the E-Sample by Person Follow-Up	67
Table A-20.A Type of Erroneous Enumerations by Relist	68
Table A-20.B Standard Errors of Type of Erroneous Enumerations by Relist	68
Table A-20.C Percent of the E-Sample by Relist	68
Table A-21.A Types of Erroneous Enumerations by ACE Regional Office	69
Table A-21.B Standard Errors of Types of Erroneous Enumerations by ACE Regional Office	70
Table A-22.A Types of Erroneous Enumerations by Domain and Tenure	71
Table A-22.B Standard Errors of Types of Erroneous Enumerations by Domain and Tenure ..	72
Table A-23: Type of Erroneous Enumeration by Post-strata	73
Table A-24: Percentage of Types of Erroneous Enumerations by Domain and Region	78
Table A-25: Standard Errors of Types of Erroneous Enumerations by Domain and Region ...	80

FIGURES

Figure 1 Weighted Number of E-Sample Returns by Week	14
Figure 2 Percent Erroneous Enumeration by Week	14
Figure 3 Type of Erroneous Enumeration by Week	15

EXECUTIVE SUMMARY

The Accuracy and Coverage Evaluation measured errors of inclusion and errors of exclusion by Census 2000. This analysis focuses on errors of inclusion, known as erroneous enumerations. This is achieved by analyzing erroneous enumerations identified in the E sample. Differences in patterns of erroneous enumerations between 1990 and 2000 were of interest to ESCAP. Reasons for the differences, such as different methodologies, changes in the population, and error in 2000 and/or 1990 were also of interest.

How did the erroneous enumeration rate compare to 1990?

In 2000, the Accuracy and Coverage Evaluation determined that 4.7 percent of the E-Sample was erroneously enumerated. In 1990, the Post-Enumeration Survey determined that 5.8 percent of the E-Sample was erroneously enumerated. **There are differences in the methodologies that make us unable to directly compare numbers.** The following changes need to be considered when comparing 2000 results to 1990 results: changes in missing data methodology, changes in timing of questionnaire mailout, differences in definition of data defined, the Duplicate Housing Unit Operation and Coverage Improvement Follow-up operation.

In 2000, 0.8 percent of the E-Sample was duplicated. In 1990, 1.6 percent of the E-Sample was duplicated. This difference can be explained by the Duplicate Housing Unit Operation performed in 2000, as well as known changes in A.C.E. methodology.

The percent insufficient information in 2000 was 1.8 percent compared to 1.2 percent in 1990. Some of this difference can be explained by the change in definition of data defined. To be considered data defined in 2000, a person had to have at least two characteristics where name counted as a characteristic. To be considered data defined in 1990, a person had to have at least two characteristics where name did not count as a characteristic.

The percent other residence, after the redistribution of people with unresolved status, was 1.4 percent in 2000 and 2.3 percent in 1990. Part of this difference may be explained by the change in coding of people who had another address, when we could not determine if it was inside or outside the search area. In other words, people who would have been coded other residence in 1990 were coded unresolved in 2000 and imputed with a high probability of erroneous enumeration. We are looking at the results of the Measurement Error Reinterview (MER) to explain this difference.

Did the erroneous enumeration rate vary across the post-strata variables as we expected?

Yes. The percent erroneous enumeration followed the same general pattern in 1990 and 2000 for Tenure, Age/Sex and Race/Hispanic Origin.

What are the types of erroneous enumerations?

By matching the people captured in the Accuracy and Coverage Evaluation to the E-Sample people, we were able to identify five types of erroneous enumerations using Accuracy and Coverage Evaluation rules and definitions: duplicates, fictitious, geocoding error, insufficient information for matching, and other residence.

Duplicates: The census counted the same person more than once. Duplicates could happen on the same form, on a different form at the same address, at a different address in the same cluster or at a different address in a surrounding block.

Fictitious: The E-Sample nonmatch was determined to be fictitious in this cluster during the Accuracy and Coverage Evaluation person follow-up interview. The person may have existed elsewhere, but the interviewer could not find anyone in the cluster who knew the person. The interviewer had to talk to at least three knowledgeable people in the cluster before a person could be considered fictitious.

Geocoding errors: If the census housing unit existed outside the A.C.E. search area, all of the people in the housing unit were erroneous enumerations.

Insufficient information for matching and follow-up: To have sufficient information for matching and follow-up, an E-Sample person had to have a complete name and at least two other characteristics. People with insufficient information for matching and follow-up were people who were data defined in the census and their name was blank, incomplete, or invalid, or they had a name and only one other characteristic.

Other residence: The Accuracy and Coverage Evaluation person follow-up interview determined that the E-Sample person was not a resident on census day because the person should have been enumerated at the other residence. The E-Sample person could have been a match or a nonmatch in before follow-up matching.

There are also people with either unresolved match or residence status. There was not enough information in Accuracy and Coverage Evaluation person follow up to determine either their match or residence status. These people had their probability of correct enumeration imputed. These people are called unresolved.

What implications do these results have on the adjustment decision?

While the results of the erroneous enumeration analysis did not bear directly on the question, the fact that **things were generally as we expected** reassures us about the quality of the A.C.E. Concerns are addressed in the Measurement Error Reinterview and ESCAP Report number 6 analyses erroneous enumerations not found during A.C.E. production.

1. BACKGROUND

1.1 Operational background

The goal of Census 2000 was to count everyone in the U.S. in their proper household. However, this did not always happen. To assess the coverage of the census, the Census Bureau undertook the Accuracy and Coverage Evaluation (A.C.E.). The A.C.E. determined whether people in the E sample, a sample of the people counted by the census in housing units, were correctly enumerated or erroneously enumerated. An erroneous enumeration occurred when the census included a person in a block cluster in error.

To determine the number of erroneous enumerations, the E-Sample people were matched to the people captured in the A.C.E. Computer and clerical matching classified E-Sample people as matched, not matched, or possibly matched. The nonmatched and possibly matched people went to A.C.E. person follow-up interviewing to determine if they were correctly or erroneously enumerated in the block cluster according to census residence rules.

- Correctly enumerated people were correctly captured in the block cluster by the census. A person that the A.C.E. person follow-up interview determined to be counted correctly by the census was a correct enumeration. An E-Sample person that matched to an A.C.E. person was also a correct enumeration.
- Erroneously enumerated person records were person records that the A.C.E. person follow-up interview determined that the census captured in error in the block cluster or search area. A person record can also be coded an erroneous enumeration during before follow-up person matching for duplicates, geocoding errors and people with insufficient information for matching and follow-up.
- If the A.C.E. person follow-up interview could not determine the person to be correctly or erroneously enumerated, the enumeration status for the E-Sample person was unresolved. Those people with unresolved enumeration status had their probability of correct enumeration imputed based on those cases that were successfully followed-up. See Childers (2001) for more details.

The rate of erroneous enumerations for a given post-stratum was used in calculating the dual system estimates. Assuming everything else is held constant, as the erroneous enumeration rate increases the dual system estimate decreases. Dual system estimates allow us to calculate undercounts, which is an important measure of the quality of the census. Understanding erroneous enumerations will help us understand the quality of the census. Knowing which variables are related to a person being erroneously enumerated will also aid in the planning for the 2010 Census.

1.2 Definitions

The E sample consisted of a sample of data defined census enumerations. To be data defined, a person had to have at least two characteristics where name counted as a characteristic.

According to A.C.E. rules there were five different types of E-Sample erroneous enumerations:

- duplicates
- other residence
- insufficient information for matching
- fictitious
- geocoding error

Duplicates: The census counted the same person more than once. Duplicates could happen on the same form, on a different form at the same address, at a different address in the same cluster or at a different address in a surrounding block.

There were cases where a person was duplicated outside of the search area, such as a college student captured at school and the parents' house. If the E-Sample person was captured in the correct place, he was coded as a correct enumeration. If the E-Sample person was captured at the incorrect place, he was coded as erroneously enumerated because he should have been enumerated at the other residence.

Other residence: The A.C.E. person follow-up interview determined that the E-Sample person was not a resident on census day because the person should have been enumerated at the other residence. The E-Sample person could have been a match or a nonmatch in before follow-up matching.

Insufficient information for matching and follow-up: To have sufficient information for matching and follow-up, an E-Sample person had to have a complete name and at least two other characteristics. People with insufficient information for matching and follow-up were people who were data defined in the census and their name was blank, incomplete, or invalid, or they had a name and only one other characteristic. However, there were 77 (unweighted) E-Sample people who were not data defined.

Fictitious: The E-Sample nonmatch was determined to be fictitious in this cluster during the A.C.E. person follow-up interview. The person may have existed elsewhere, but the interviewer could not find anyone in the cluster who knew the person. The interviewer had to talk to at least three knowledgeable people in the cluster before a person could be considered fictitious.

Geocoding errors: If the census housing unit existed outside the A.C.E. search area, all of the people in the housing unit were erroneous enumerations.

There were also people with either unresolved match or residence status. There was not enough information in A.C.E. person follow up to determine either their match or residence status.

These people had their probability of correct enumeration imputed. In section 4.1, I combined all of the people with unresolved status into a category called unresolved. In section 4.2, I put the people with unresolved status into the following categories based on how they were imputed:

- other residence
- fictitious
- geocoding error

It should be noted that a different definition of what is considered erroneous would lead to a different erroneous enumeration rate.

2. METHODS

The erroneous enumeration rate was the weighted number of people in the E sample that were erroneously enumerated divided by the total weighted number of people in the E sample. To determine the number of erroneously enumerated people, I used the probability of erroneous enumeration (one minus the probability of correct enumeration). Rates for the different types of erroneous enumerations were calculated similarly, with the numerator being the number of that type of erroneous enumeration and the denominator being the total number of people in the E sample. See Appendix B for more details on calculating erroneous enumeration rates.

To take into consideration the complex survey design, I used stratified Jackknife method and VPLX to compute the standard errors. All hypothesis testing were two-tailed at the 0.10 significance level. Bonferroni's adjustment was used for multiple comparisons. This analysis includes the 50 states and the District of Columbia.

3. LIMITS

This paper does not fully explore the interactions between variables (see O.12 Analysis of Nonmatches and Erroneous Enumerations Using Logistic Regression). This paper has a specific focus on erroneous enumerations. Other issues or errors beyond the scope of this paper include:

- P-sample nonmatches, i.e., errors of exclusion (see O.6 Accuracy and Coverage Evaluation Persons Not Matched in Census 2000)
- Matching error (see N.14 Evaluation of Matching Error)
- Response error in reported characteristics
- Imputation error in correcting ambiguity or inconsistency in census data
- Error due to whole-household nonresponse
- Correlation bias or lack of independence between census and A.C.E. enumerations resulting in understated or overstated erroneous enumeration rates
- Analysis of census people out of the scope of the A.C.E. (for example Group Quarters people)

4. RESULTS

4.1 The Different Types of Erroneous Enumerations

The breakdown of types of erroneous enumerations can be seen in Table 1. The number given include the final E-Sample person weight that reflects the probability of selection in all stages of sampling including Targeted Extended Search sampling, noninterview adjustment and weight trimming. The number reflects a ratio adjustment of the E-Sample Universe to the weighted total E-Sample estimate (see Appendix B for details).

Table 1 Types Erroneous Enumerations in Census 2000 (standard errors)		
Type	Number	Percent
Duplicate	2,014,675 (67,435)	0.8 (0.03)
Fictitious	708,285 (54,608)	0.3 (0.02)
Geocoding Error	636,053 (110,472)	0.2 (0.04)
Other Residence	2,744,138 (67,376)	1.0 (0.03)
Insufficient Information	4,781,418 (100,476)	1.8 (0.02)
Unresolved	1,645,321 (59,706)	0.6 (0.02)
Total	12,529,889 (213,020)	4.7 (0.07)

The following sections describe the types of erroneous enumerations for the following variables: race/Hispanic origin, age/sex, tenure, return rate, place size/type of enumeration area, region, response method, imputation, housing unit match status and mail return date. For each of these variables, there is a table of the types of erroneous enumerations, a table of standard errors, a table of significant differences and a table with the percent of E sample for the variable in Appendix A. Appendix A also contains tables that are not mentioned in the text of this paper.

4.1.1 Race/Hispanic origin

For post-stratification purposes, there were seven race/Hispanic origin groups: American Indian on reservation, American Indian off reservation, Hispanic, Non-Hispanic black, Native Hawaiian or Pacific Islander, Non-Hispanic Asian and Non-Hispanic white. Haines (2001) explains how

multi-racial people were placed into categories. See Table 2 for the percent erroneous enumeration.

The following is a brief explanation of which race groups were significantly different with respect to total erroneous enumeration rate:

- **American Indians on reservations** had a lower rate than American Indians off reservations, non-Hispanic blacks and Hispanics.
- **American Indians off reservations** had a higher rate than American Indians on reservations and non-Hispanic whites.
- **Hispanics** had a higher rate than American Indians on reservations and non-Hispanic whites. They had a lower rate than non-Hispanic blacks.
- **Non-Hispanic blacks** had a higher rate than American Indians on reservations, non-Hispanic Asians, non-Hispanic whites and Hispanics.
- **Hawaiians and Pacific Islanders** had a higher rate than non-Hispanic whites.
- **Non-Hispanic Asians** had a higher rate than non-Hispanic whites and a lower rate than non-Hispanic blacks.
- **Non-Hispanic whites** had a lower rate than all other categories except American Indians on reservations.

The different type of erroneous enumerations are difficult to interpret for race/Hispanic origin. The results given here are broad generalizations. Non-Hispanic blacks and Hispanics had high duplicate rates, 1.2 percent and 1.1 percent respectively. Non-Hispanic blacks also have a high fictitious rate, 0.7 percent. American Indians on reservations had a low rate of insufficient information, 0.8 percent. However, American Indians on reservations had a high other residence rate, 1.7 percent. Non-Hispanic whites had a low unresolved rate, 0.4 percent. See Tables A-1.A, B, C, and D for the percentage in each type of erroneous enumeration by race/Hispanic origin, standard errors, details about significant differences and percent of the E sample.

Table 2 Percent Erroneous Enumerations by Domain (standard errors)

1. American Indian on reservation	4.2 (0.34)
2. American Indian off reservation	6.0 (0.56)
3. Hispanic	5.5 (0.18)
4. Non-Hispanic black	7.3 (0.21)
5. Hawaiian or Pacific Islander	7.0 (1.00)
6. Non-Hispanic Asian	5.4 (0.32)
7. Non-Hispanic white	4.1 (0.07)
Total	4.7 (0.07)

4.1.2 Age/Sex

There were seven age/sex categories used for post-stratification: 0-17 years of age, 18-29 male, 18-29 female, 30-49 male, 30-49 female, 50+ male, and 50+ female. Past evidence showed that 18-29 year old people are difficult to count. They tend to be more mobile than other age categories. See Table 3 for the percent of erroneous enumerations broken down by age/sex category. The following explains which values were significantly different with respect of erroneous enumeration rates:

- **0-17 years of age** had a lower rate than all other categories except 30-49 females.
- **18-29 males** had a higher rate than all other categories.
- **18-29 females** had a higher rate than 0-17 years of age, 30-49 males, 30-49 females, 50+ males and 50+ females. They had a lower rate than 18-29 males.
- **30-49 males** had a higher rate than 0-17 years of age and 30-49 females. They had a lower rate than 18-29 males and 18-29 females.
- **30-49 females** had a lower rate than all other categories except 0-17 year of age.
- **50+ males** had a higher rate than 0-17 years of age and 30-49 females. They had a lower rate than 18-29 males and 18-29 females.
- **50+ females** had a higher rate than 0-17 years of age and 30-49 females. They had a lower rate than 18-29 males and 18-29 females.

Although I used the post-stratification categories to analyze these data, it is interesting to note that 0-17 males did not differ from 0-17 females in total erroneous enumeration rate or by any of the different types of erroneous enumerations. The people 0-17 years of age had a significantly lower duplicate rate (0.6 percent) than all of the other age/sex categories. Males and female 18-29 had high other residence and unresolved rates. See Tables A-2.A, B, C, and D for more detailed information about age/sex.

**Table 3 Percent Erroneous Enumerations
by Age/Sex (standard errors)**

0-17	4.1 (0.09)
18-29 Male	7.1 (0.16)
18-29 Female	6.4 (0.15)
30-49 Male	4.8 (0.11)
30-49 Female	4.0 (0.09)
50+ Male	4.7 (0.11)
50+ Female	4.5 (0.10)
Total	4.7 (0.07)

4.1.3 Tenure (Owner vs. Non-Owner)

Past evidence suggested that owners have a lower erroneous enumeration rate than non-owners. Owners tend to live in the same place longer and have more connections to the community than non-owners. In Census 2000, owners had a lower erroneous enumeration rate than non-owners (see Table 4).

The owners had a lower duplicate rate, fictitious rate, insufficient information rate, other residence rate and unresolved rate than non-owners. The only rate in which owners and non-owners did not differ was the geocoding error rate. See Tables A-3.A, B, C, and D for more detailed information about tenure.

Table 4 Percent Erroneous Enumerations by Tenure (standard errors)		• Owners had a significantly lower erroneous enumeration rate than non-owners
Owner	3.6 (0.08)	
Non-Owner	7.3 (0.13)	
Total	4.7 (0.07)	

4.1.4 Return rate indicator

Return rates are an important indicator of public cooperation with the census. Tract-level return rates are calculated for each tract with mailback enumeration areas. Areas with high return rates are expected to have lower rates of erroneous enumerations than areas with low return rates.

Return rate is an A.C.E. post-stratification variable for the Non-Hispanic White or “Some other race,” Non-Hispanic Black, and Hispanic domains. Therefore, E-Sample persons in these three race/Hispanic origin domains are affiliated with a high or low return rate indicator value. E-sample persons in all other race/Hispanic origin domains are assigned a return rate indicator value of “Not Applicable” since they are not post-stratified by return rate. See Kostanich (2001) for details on return rate calculations and the high/low designation.

Table 5 shows that E-Sample persons associated with high return rate indicator values have lower erroneous enumeration rates than both E-Sample people with low return rate indicator values and E-Sample persons who were not post-stratified by the return rate variable. E-Sample persons affiliated with high return rate indicator values also have lower duplicate rates, fictitious rates, and unresolved rates than E-Sample persons in the other two return rate categories. Finally, E-Sample persons associated with high return rate indicator values have lower insufficient information rates and other residence rates than people associated with low return rate indicator values. See Tables A-4.A, B, C, D for more detailed information about return rate.

Table 5 Percent Erroneous Enumerations by Return Rate (standard errors)

High	4.2 (0.08)
Low	6.1 (0.14)
Not Applicable	5.5 (0.27)
Total	4.7 (0.07)

- Areas with high return rates had a lower erroneous enumeration rate than areas with low return rates and areas that were not mail return areas

4.1.5 Place size and type of enumeration

Metropolitan Statistical Areas (MSA) were broken down into four categories: large, medium, small and non-MSA. These MSA categories were combined with information about how people got their form, type of enumeration area (TEA). Type of enumeration areas were broken down into two categories: mail out/mail back (MO/MB) and all other TEAs.

Table 6 shows that large MSA, mail out/mail back areas had a higher rate of erroneous enumeration than all of the other categories. Large MSA mail out/mail back tended to have high duplicate rate, fictitious rates and insufficient information rates than the other categories. See Tables A-5.A, B, C, and D for more detailed information about place size and type of enumeration area.

Table 6 Percent Erroneous Enumerations by Metropolitan Statistical Area/Type of Enumeration Area (standard errors)

Large MSA MO/MB	5.2 (0.16)
Medium MSA MO/MB	4.5 (0.12)
Small MSA & Non-MSA MO/MB	4.4 (0.15)
All Other TEAs	4.6 (0.13)
Total	4.7 (0.07)

- Large MSAs, mail out/mail back areas had a higher erroneous enumeration rate than medium MSA, mail out/mail back areas; small MSA and non-MSA, mail out/mail back areas; and all other TEAs.

4.1.6 Region

The Census Bureau divided the country into four regions: Northeast, Midwest, South and West. Midwest had a lower rate of erroneous enumerations than the other regions (see Table 7). The Midwest tended to have lower rates of insufficient information, other residence and unresolved. The Northeast's duplicate rate was higher than the other regions. See Tables A-6.A, B, C, and D for more detailed information about region.

**Table 7 Percent Erroneous
Enumerations by Region (standard
errors)**

Northeast	5.0 (0.16)
Midwest	3.8 (0.13)
South	5.1 (0.14)
West	4.8 (0.15)
Total	4.7 (0.07)

- **The Midwest had a lower erroneous enumeration rate than the Northeast, the South and the West.**

4.1.7 Response method

Most households were self-reporting (mail returns and internet returns). However, 23.7 percent of people in the E sample gave their information to an enumerator. An enumerator visited housing units in areas without reliable mail delivery, areas with a high percentage of people who used post-office boxes and people who did not mail back their census form. The enumerators tried to get an interview with a household member. Sometimes this was not possible, so the enumerator had to get a proxy interview with someone outside the household. Of the enumerator filled returns, 11.3 percent were with a proxy respondent. In this section, I compare self-reporting, enumerator returns completed by a household member and enumerator returns completed by a proxy respondent.

The following explains which values had significantly different erroneous enumeration rates (see Table 8):

- **Self-reporting responses** had a lower erroneous enumeration rate than both the enumerator returns completed by a household member and enumerator returns completed by a proxy respondent.
- **Enumerator returns completed by a household member** had a higher erroneous enumeration rate than self-reporting responses and a lower erroneous enumeration rate than the enumerator returns completed by a proxy respondent.
- **Enumerator returns completed by a proxy respondent** had a higher erroneous enumeration rate than enumerator returns completed by a household member and self-reporting responses.

The following explains which values had significantly different duplicate rates, insufficient information rates, other residence rates and unresolved rates (see Tables A-7.A, B, C, and D for more detailed information about response method):

- **Self-reporting responses** had a lower rate than both the enumerator returns completed by a household member and enumerator returns completed by a proxy respondent.
- **Enumerator returns completed by a household member** had a higher rate than self-reporting responses and a lower rate than the enumerator returns completed by a proxy respondent.
- **Enumerator returns completed by a proxy respondent** had a higher rate than enumerator returns completed by a household member and self-reporting responses.

Self-reporting people had a lower fictitious rate than both the enumerator returns completed by a household member and enumerator returns completed by a proxy respondent, but enumerator returns completed by a household member and enumerator returns completed by a proxy respondent did not differ from each other.

Table 8 Percent Erroneous Enumerations by Response Method (standard errors)	
Self-reporting	2.9 (0.06)
Enumerator returns	10.6 (0.19)
Household member	7.1 (0.16)
Proxy	37.2 (0.81)
Total	4.7 (0.07)

4.1.9 Characteristic Imputation

Some people did not answer all of the census questions. When this happened, we imputed missing characteristic for the person in the census. There were also cases where the data were edited through consistency edits. I looked at the erroneous enumeration rates of people who had no imputation and no data edits verses those with some imputations or some data edits.

I found that people with some imputations or some data edits had a higher erroneous enumeration rate than those people with no imputations and no data edits (see Table 9). In the E sample, 13.0 percent of the people had some imputation or some data edits. People with some imputations or some data edits had a higher duplicate rate, fictitious rate, insufficient information rate, other residence rate and unresolved rate. See Tables A-8.A, B, C, and D for more information about imputation.

Table 9 Percent Erroneous Enumerations by Characteristic Imputation (standard errors)	
	Total
No Characteristic Imputation and No Data Edits	3.3 (0.07)
Some Characteristic Imputation or Data Edits	14.5 (0.25)
Total	4.7 (0.07)

4.1.10 Housing unit's match status

Before person matching occurred, the Census Bureau conducted a housing unit match. There were four possible outcomes for the housing unit matching: correctly enumerated, erroneously enumerated, unresolved and no matching. Housing units with no matching were added to the inventory of census housing units since January 2000.

- **People in correctly enumerated housing units** had a significantly lower erroneous enumeration rate than people in erroneously enumerated housing units, housing units with

unresolved enumeration status and housing units where no matching took place (see Table 10).

- **People in erroneously enumerated housing units, housing units with unresolved enumeration status and housing units where no matching took place** did not differ in their erroneous enumeration rates.

This shows that a person's chances of being erroneously enumerated were related to how well the census did in locating their housing unit. Of the E sample, 92.4 percent of the people lived in correctly enumerated housing units, 4.6 percent lived in erroneously enumerated housing units, 0.1 percent lived in housing units with unresolved enumeration status and 2.9 percent lived in housing units that did not go through housing unit matching. See Tables A-9.A, B, C, and D for more detailed information on housing unit match status.

**Table 10 Percent Erroneous Enumerations
by Housing Unit's Match Status (standard
errors)**

Correctly Enumerated	4.0 (0.06)
Erroneous Enumerated	14.5 (0.75)
Unresolved Enumeration Status	12.9 (2.71)
No Matching	11.9 (1.05)
Total	4.7 (0.07)

It is also interesting to consider people by their housing unit's initial address status. The housing unit could be: matched in the housing unit phase, not matched in the housing unit phase, added to the DMAF after the housing unit phase, or a conflicting household. A conflicting household was a household where the A.C.E. housing unit matched the census housing unit and both contain whole households of nonmatched people. This has been called a Smith/Jones household.

Table 11 shows the erroneous enumeration rates for the people in the different types of housing units. The following explains which values are significantly different:

- **Housing unit matched** had a lower rate than housing unit not matched, housing unit added and conflicting households.
- **Housing unit not matched** had a higher rate than housing unit matched and lower than conflicting household.
- **Housing unit added** had a higher rate than housing unit matched and lower than conflicting household.
- **Conflicting household** had a higher rate than all other categories.

Duplicate, fictitious, other residence and unresolved rates followed the same pattern. This shows that housing units captured in both the census and A.C.E. had high erroneous enumeration rates

when the census and A.C.E. captured conflicting households. See Tables A-10.A, B, C, and D for more detailed information on initial address status.

Table 11 Type of Erroneous Enumeration by Initial Address Status (standard errors)	
HU matched	3.4 (0.05)
HU not matched	13.0 (0.57)
HU added to DMAF after HU phase of A.C.E.	11.9 (1.05)
Conflicting households	36.5 (1.09)
Total	4.7 (0.07)

4.1.11 Mail Return Date

For census mail returns, we had information about the date the mail return was checked in to the data processing center. Figure 1 shows the weighted number of E-Sample mail returns checked in by week. Note that the entire months of July and August are each combined. This was due to the fact that check-in only occurred on a couple of days each of those months. Figure 1 shows that most of the mail returns were checked-in between March 3, 2000 and April 21, 2000. Figure 2 shows the percent erroneous enumeration for mail returns by their check-in date. Figure 2 shows that the percent erroneous enumeration generally increases over time. Figure 3 shows the percent of the various types of erroneous enumerations for mail returns by their check-in date. Figure 3 show the same general upward trend for all of the types of erroneous enumerations, except geocoding error. This is the same pattern seen in Figure 2. Note that Figure 3 excludes August.

Figure 1 Weighted Number of E-Sample Returns by Week

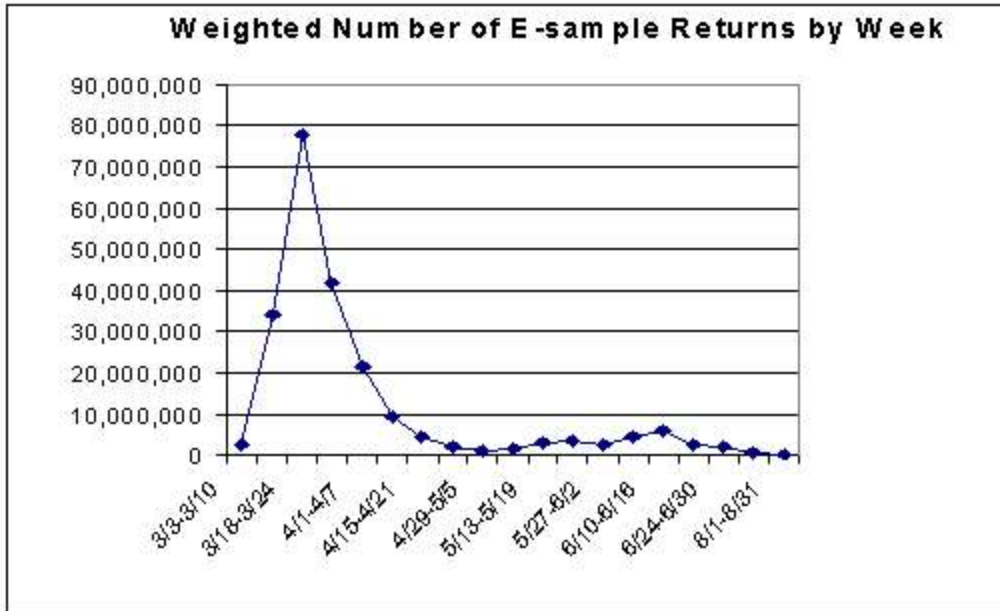


Figure 2

Percent Erroneous Enumeration by Week

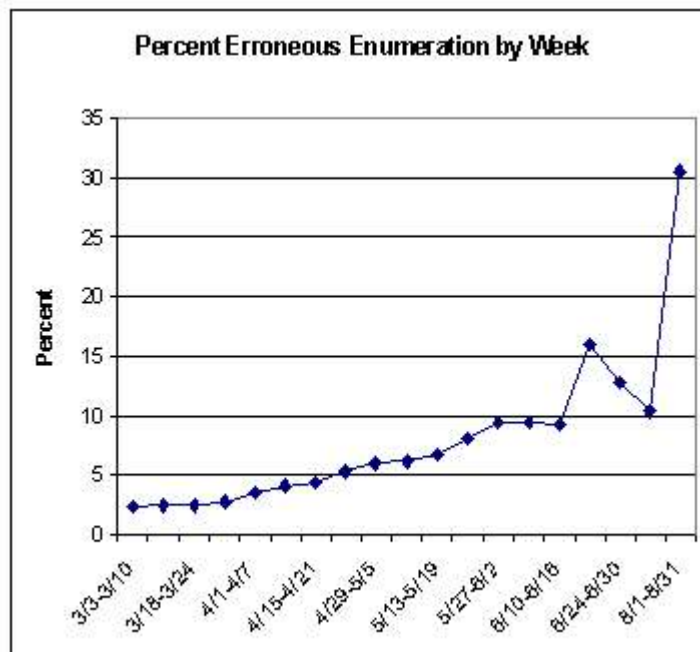
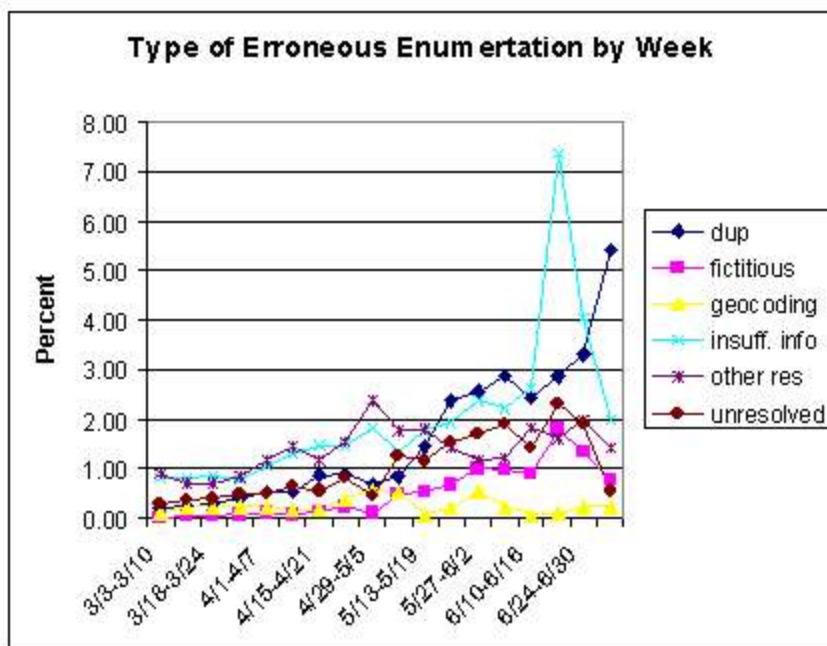


Figure 3 Type of Erroneous Enumeration by Week



4.1.12 Other variables

Appendix A also contains detailed tables of types of erroneous enumerations, standard errors, results of significance testing and percent of the E-Sample for the following variables:

- Number of units at a basic street address (Tables A-11.A, B, C, D)
- Whole/partial match code (Tables A-12.A, B, C, D)
- Form length (Tables A-13.A, B, C, D)
- Non-response follow up (Tables A-14.A, B, C, D)
- Coverage edit follow up (Tables A-15.A, B, C, D)
- Coverage improvement follow up (Tables A-16.A, B, C, D)
- Capture method (Tables A-17.A, B, C, D)
- Data Capture Center (Tables A-18.A, B, C, D)
- A.C.E. person follow-up (Tables A-19.A, B, C)
- A.C.E. relisted cluster (Tables A-20.A, B, C)
- A.C.E. regional office (Tables A-21.A, B)
- Race/Hispanic origin and tenure (Tables A-22.A, B)
- Post-strata variables (Tables A-23)

4.2 What is the difference in erroneous enumeration between 1990 and 2000

4.2.1 Changes between 1990 and 2000 that affect erroneous enumeration rates

There were some differences between the 1990 Post-Enumeration Survey (PES) and the 2000 A.C.E. that should be considered when comparing erroneous enumeration rates.

- The E-Sample universe changed between 1990 and 2000. Non-institutional, non-military group quarters were included in the E-Sample universe in 1990, but were excluded from the universe in 2000.

If the unit was a(n)...	in 1990 it was...	in 2000 it was...
housing units	in the universe	in the universe
non-institutional, non-military group quarters	in the universe	not in the universe
institutional group quarters	not in the universe	not in the universe
military group quarters	not in the universe	not in the universe
unit in remote Alaska	not in the universe	not in the universe

- To be included in the E-Sample universe, a person had to be data defined. There was a change in the definition of data defined between 1990 and 2000. In 1990, people were data defined if they had two characteristics where name did not count as a characteristic. In 2000, names were captured for the first time. People were data defined if they had two characteristics where name counted as a characteristic¹. In 1990, the first person on the form automatically had the relationship of reference person that counted as a characteristic. In 2000, the relationship of the reference person did not count as a characteristic.
- Unclassified means the number of people in housing unit had to be estimated. The number of unclassifieds was larger in 2000 than it was in 1990.
- In 2000, an operation called the Duplicate Housing Unit Operation (Nash, November, 2000) was developed to get rid of some of the census duplication. The operation eliminated duplicates before they went to matching. In 1990, there was no such operation. This means that the duplicate rate should be lower in 2000, because some duplicates were eliminated before matching. The duplicate operation had a larger search area for duplicates than A.C.E. person matching. The removal of duplicates outside the A.C.E. search area should lower the other residence rate.
- In 2000, the Duplicate Housing Unit Operation identified potential duplicates. These people were excluded from the E Sample. Nearly 2.2 million people were later re-instated. This means that the A.C.E. did not directly measure the erroneous enumeration rate for nearly 2.2 million people that were in the E sample universe. In 1990, a smaller

¹ In 2000, there were 77 people included in the E sample who were not data defined. They were coded insufficient information for matching.

group of people were imputed as late census data. The reinstated people were treated the same as whole person imputations in the dual system estimator.

- There was a change in the search area for duplicates. In 1990, rural areas had a search area of two rings of surrounding block for duplicates. In 2000, the search area was limited to the block cluster. However, in 2000, there was a targeted extended search that expanded the search area for duplicates to the first ring of surrounding blocks for clusters likely to benefit from an expanded search area. For example, if there was a duplicate in the sample block cluster and one in the first ring of surrounding blocks, this case would be handled differently in 1990 and 2000. In 1990, the E-Sample person duplicated to the surrounding block was given half an erroneous enumeration. In 2000, the E-Sample person was searched for in the surrounding blocks only when there was evidence of geocoding error. If the E-Sample person was correctly geocoded, there was no duplicate search. If the E-Sample person was incorrectly geocoded and should have been in the first ring of surrounding blocks, we expanded the search area to the surrounding block and searched for duplicates. The E-Sample person would have been a full erroneous enumeration, if duplicated, otherwise the person was a correct enumeration.
- The targeted extended search caused a difference in rural areas. In 1990, a person duplicated to the second ring of surrounding blocks would have been considered half a correct enumeration. In 2000, the E-Sample person would be a correct enumeration if the person should have been counted at the sample address and an erroneous enumeration if the person should have been counted at the other address.
- There was a change in the way the probability of correct enumeration was computed for duplicates. This is due to the fact that, in 2000, we had more information about the housing unit geography.
- In 1990, the Post-Enumeration Survey included people in noninstitutional, nonmilitary group quarters. When we compare erroneous enumeration rates between 1990 and 2000, we are assuming that the people in these group quarters had the same pattern of erroneous enumerations as those people in housing units. However, people in group quarters may have a different erroneous enumeration rate than people in housing units.
- The instructions about who to include on the census form changed between 1990 and 2000.
- The order in which the erroneous enumerations were coded has changed between 1990 and 2000. See Childers (January, 2001) for coding in Census 2000. See Childers (September, 2001) for coding in the 1990 Census.
- In 1990, probabilities of correct enumeration were assigned to people with unresolved status using a logistic regression model. In 2000, probabilities of correct enumeration were assigned to people with unresolved status using Imputation Cell Estimation. This

change could have impacted the correct enumeration probabilities; however, it should not have significantly change the percent erroneous enumeration due to unresolved status.

- There was a change in the missing data procedures that caused the percent erroneous enumeration due to unresolved status to increase. The procedure used information from the A.C.E. person follow-up questionnaire to create two new imputation cells. These cells were assigned a lower probability of correct enumeration than the remaining people with unresolved status (Cantwell, March, 2001).
- In 2000, housing units that were vacant or deleted in nonresponse follow-up and deleted during questionnaire delivery were not sent to Census Improvement Follow-up (Hogan, June, 2000). In 1990, the Vacant/Delete/Movers Check sent housing units classified as vacant or delete during List/Enumerate or Nonresponse Follow-up to field follow-up. The field follow-up completed questionnaires for housing units that changed status (1990 CPH-E-3). People in these types of housing units would have been included in the E-Sample universe in 1990, but not 2000. These people may have had a higher erroneous enumeration rate than other people.
- The Census Bureau mailed out the Census questionnaires earlier in 2000 than in 1990. People who filled out their census form a long time before census day may have had a higher chance of moving before Census Day.
- There was a slight change in the coding of other residence and usual home elsewhere. This affected people counted at a special place who could claim usual home elsewhere: people captured on military forms, shipboard people, people captured on individual census questionnaires, and people captured on individual census reports.

If a person captured at a special place...	in 1990 it was coded...	in 2000 it was coded...
did not claim usual home elsewhere	correct enumeration	correct enumeration
claimed usual home elsewhere	other residence	correct enumeration

This means that some people who would have been erroneously enumerated due to other residence in the 1990 Census were considered correct enumerations due to usual home elsewhere in Census 2000. This affected about 14,000 people.

- There was also a coding change that affected the other residence rates and the unresolved rate:

If a person had another address...	in 1990 it was coded...	in 2000 it was coded...
and the address was outside the search area	other residence	other residence
and we could not determine if the address was inside or outside the search area	other residence	unresolved
and no address was given	other residence	unresolved

A person was erroneously enumerated when they should have been counted at another address unless that other address was in the search area. In 1990, we coded these people as erroneous unless we knew the other address was in the search area. In 2000 we coded these people as unresolved when we were not certain. In 2000, the imputation procedures took this into consideration. This affected about 1.1 million people.

- There was a change in the search area when considering nonmatches. In 1990, clerks checked for matches and duplicates in the surrounding blocks for all clusters. In 2000, there was a targeted extended search that selected clusters which were likely to benefit from a surrounding block search. See analysis project N.17 - Target Extended Search Analysis for the effect of targeted extended search on a person's probability of being erroneously enumerated. Table 12 shows the effect of targeted extended search on erroneous enumeration rates.

Table 12 Percent Erroneous Enumerations Before and After TES

	pre- TES	post- TES	Difference
Duplicate	0.72	0.76	-0.04
Fictitious	0.26	0.27	-0.01
Geocoding Error	3.30	0.24	3.06
Other Residence	0.99	1.03	-0.04
Insufficient Information	1.80	1.80	0.00
Unresolved	0.56	0.62	-0.06
Total	7.63	4.72	2.91

4.2.2 A comparison of erroneous enumeration rates between 1990 and 2000

Table 3 shows the 1990 and 2000 erroneous enumeration rates. A direct comparison of rates is limited by the points given in section 4.2.1.

**Table 13 Comparison Between 1990
and 2000 of the Percent Erroneous
Enumerations**

	2000	1990
Duplicate	0.76	1.62
Fictitious	0.27	0.15
Geocoding Error	0.24	0.34
Other Residence	1.03	2.18
Insufficient Information	1.80	1.19
Unresolved	0.62	0.26
Total	4.72	5.74

Note: 1990 data are from Childers (September, 2001) and related to the PES universe

4.2.3 Reclassification of people with unresolved status

People with unresolved status were people for whom there was not enough information in the A.C.E. person follow up to determine their match or enumeration status. However, we gained some information about the enumeration status of these people by looking at their follow-up forms. Based on follow-up information, we could classify them into the following categories: fictitious, geocoding error, other residence. This information was also used to impute the person's probability of correct enumeration during the missing data operation (Cantwell, 2001). See Appendix B for details on the redistribution of people with unresolved status.

Table 4 shows how the rates of the various types of erroneous enumerations changed when the people with unresolved status were incorporated into the different rates. The fictitious rate, geocoding error rate and other residence rate increased.

Table 14: Comparison Between 1990 and 2000 of the Percent Erroneous Enumerations After Re-classifying the Unresolved People

	2000	1990
Duplicate	0.76	1.66
Fictitious	0.50	0.22
Geocoding Error	0.25	0.38
Other Residence	1.41	2.29
Insufficient Information	1.80	1.19
Total	4.72	5.74

Note: 1990 data are based on Childers (September, 2001) and related to the PES universe

SUMMARY

Highlights of the different types of erroneous enumerations are given below:

- Duplicate rates were high in large cities in the Northeast, 1.5 percent. Duplicate rates were also high for people in housing units that were not correctly enumerated, 4.1 percent.
- Conflicting households, households where the census captured one family and the A.C.E. captured another family, had high fictitious rates, 8.8 percent.
- In all variables I analyzed, erroneous enumerations due to geocoding error were insignificant in all tests. Targeted Extended Search procedures reduced the effects of erroneous enumerations due to geocoding error by allowing correct enumerations in the surrounding blocks.
- Insufficient information was the highest in enumerator filled returns (4.7 percent), especially if the respondent was a proxy (27.4 percent). American Indians on reservations had low rates of insufficient information, 0.8 percent. Of the people with insufficient information, 30.3 percent were called insufficient information because they had an invalid name, such as Donald Duck.
- The Midwest had low rates of other residences, 0.8 percent, and American Indians on reservations had high rates of other residences, 1.7 percent.

- There was a very low unresolved rate, 0.2 percent, in the Boston Regional Office which covers Maine, Vermont, New Hampshire, Massachusetts, Rhode Island, Connecticut and up-state New York. People 18-29 years of age had a high unresolved rate, 1.3 percent.

References

For a complete analysis of the different types of erroneous enumerations see the following analysis projects:

- *Duplicates*: O.16 - Person Duplication in the 2000 Census. This is an analysis of person duplication captured by the A.C.E.
- *Fictitious*: N.10 - Evaluation of Falsification in A.C.E. This analysis project briefly looks at people coded fictitious in the A.C.E. and compares it with people coded fictitious during the Evaluation follow-up.
- *Geocoding error*: O.9 - Geocoding Error Analysis. This will also look at geocoding error in the first ring of surrounding blocks.
- *Insufficient information for matching*: O.23 - Analysis of Insufficient Information for Person Matching.
- *Other residence*: N.8 - Analysis of the Census Residence Questions used in A.C.E. This analysis project is not dedicated specifically to other residence. It's main focus is residence question in the A.C.E. person interview.
- *Unresolved*: N.13 - Analysis of Unresolved Codes in Person Matching.

1990 CPH-E-3, "Programs to Improve Coverage in the 1990 Census", 1990 Census of Population and Housing Evaluation and Research Reports, U.S. Department of Commerce, Economics and Statistics Administration, BUREAU OF THE CENSUS.

Cantwell, P and Childers, D. (March 2001), "Accuracy and Coverage Evaluation Survey: A Change to the Imputation Cells to Address Unresolved Resident and Enumeration Status", DSSD Census 2000 Procedures and Operations Memorandum Series, Chapter Q-44.

Childers, D. (January 2001), "Accuracy and Coverage Evaluation: The Design Document", DSSD Census 2000 Procedures and Operations Memorandum Series, Chapter S-DT-01.

Childers, D. (September, 2001), "1990 E-Sample Documentation", Internal Memorandum.

Cromar, R. (June, 2000), "Accuracy and Coverage Evaluation Survey: Additions to the Sample Design File Version 5", DSSD Census 2000 Procedures and Operations Memorandum Series, Chapter R-34.

Haines, D. (March, 2001), "Accuracy and Coverage Evaluation Survey: Computer Specifications for Person Dual System Estimation (U.S.) -Re-issue of Q-37", DSSD Census 2000 Procedures and Operations Memorandum Series, Chapter Q-48.

Hogan, H. (September 1993), "The 1990 Post-Enumeration Survey: Operations and Results", Journal of the American Statistical Association, Vol 88, No 423.

Hogan, H. (June, 2000), “Revision Specification of the Coverage Improvement Followup Universe for Census 2000”, “DSSD Census 2000 Procedures and Operations Memorandum Series, Chapter CC-3Revised#4.

Jones, C. (December, 1999), “Requirements for Person Matching Review and Coding System (PerMaRCS) File Transfers”, DSSD Census 2000 Procedures and Operations Memorandum Series, Chapter S-CO-03.

Nash, F. (November, 2000), “Overview of the Duplicate Housing Unit Operations”.

Philipp, D. (February, 2001), “Hundred Percent Census Unedited File (HCUF)”, 2000 Decennial Documentation.

Philipp, D. (April, 2001), “Hundred Percent Census Edited File (HCEF)”, 2000 Decennial Documentation.

Appendix A
Detailed Tables of Types of Erroneous Enumerations, Standard Errors, Results of
Significance Testing and Percent of the E-Sample

Race/Hispanic Origin

Table A-1.A Types of Erroneous Enumerations by Domain							
	Duplicate	Fictitious	Geocoding Error	Insuff. Info	Other Resid.	Unresolved	Total
1. AI On Reservation	0.7	0.1	0.1	0.8	1.7	0.7	4.2
2. AI Off Reservation	0.7	0.3	0.1	2.7	1.2	1.0	6.0
3. Hispanic	1.1	0.5	0.3	1.7	0.9	1.1	5.5
4. Non-Hispanic Black	1.2	0.7	0.3	3.0	1.1	1.1	7.3
5. Native Hawaiian or Pacific Islanders	0.8	0.5	0.0	4.4	0.5	0.8	7.0
6. Non-Hispanic Asian	1.0	0.4	0.3	1.8	1.1	0.8	5.4
7. Non-Hispanic White	0.6	0.2	0.2	1.6	1.0	0.4	4.1
Total	0.8	0.3	0.2	1.8	1.0	0.6	4.7

Table A-1.B Standard Errors of Types of Erroneous Enumerations by Domain

	Duplicate	Fictitious	Geocoding Error	Insuff. Info	Other Resid.	Unresolved	Total
1. AI On Reservation	0.14	0.03	0.09	0.18	0.25	0.13	0.34
2. AI Off Reservation	0.17	0.10	0.08	0.46	0.23	0.22	0.56
3. Hispanic	0.07	0.06	0.08	0.08	0.05	0.07	0.18
4. Non- Hispanic Black	0.08	0.08	0.11	0.12	0.06	0.08	0.21
5. Native Hawaiian or Pacific Islanders	0.20	0.23	0.01	0.88	0.19	0.21	1.00
6. Non- Hispanic Asian	0.18	0.08	0.17	0.16	0.10	0.09	0.32
7. Non- Hispanic White	0.03	0.02	0.04	0.04	0.03	0.02	0.07
Total	0.03	0.02	0.04	0.03	0.02	0.02	0.07

Table A-1.C Significant Differences among Domain

Rates	Which values are significant (t >2.807)
Total Erroneous Enum. Rate	<p>AI on reservation differed from AI off reservation, Hispanic and Non-Hispanic Black.</p> <p>AI off reservation differed from AI on reservation and Non-Hispanic White.</p> <p>Hispanic differed from AI on reservation, Non-Hispanic Black and Non-Hispanic White.</p> <p>Non-Hispanic Black differed from AI on reservation, Hispanic, Non-Hispanic Asian and Non-Hispanic White.</p> <p>Native Hawaiian or Pacific Islander differed from Non-Hispanic Whites.</p> <p>Non-Hispanic Asian differed from Non-Hispanic Black and Non-Hispanic White.</p> <p>Non-Hispanic White differed from all other categories except AI on reservation.</p>
Duplicate Rate	<p>AI off reservation had a lower rate than non-Hispanic blacks.</p> <p>Hispanics had a higher rate than non-Hispanic whites.</p> <p>Non-Hispanic blacks had a higher rate than AI off reservation and non-Hispanic whites.</p> <p>Non-Hispanic white had a lower rate than non-Hispanic blacks and Hispanics.</p>
Fictitious Rate	<p>AI on reservation differed from Hispanic, Non-Hispanic Black and Non-Hispanic Asian.</p> <p>AI off reservation differed from Non-Hispanic Black.</p> <p>Hispanic differed from AI on reservation and Non-Hispanic White.</p> <p>Non-Hispanic Black differed from AI on reservation, AI off reservation and Non-Hispanic White.</p> <p>Native Hawaiian or Pacific Islander did not differ from the other categories.</p> <p>Non-Hispanic Asian differed from AI on reservation and Non-Hispanic White.</p> <p>Non-Hispanic White differed from Hispanic, Non-Hispanic Black and Non-Hispanic Asian.</p>
Geocoding Rate	<p>Native Hawaiian or Pacific Islander differed from Hispanic and Non-Hispanic White.</p>

Insufficient Info Rate	<p>AI on reservation are lower than all other categories.</p> <p>AI off reservation differed from AI on reservation.</p> <p>Hispanic differed from AI on reservation, Non-Hispanic Black and Native Hawaiian or Pacific Islander.</p> <p>Non-Hispanic Black differed from AI on reservation, Hispanic, Non-Hispanic Asian and Non-Hispanic White.</p> <p>Native Hawaiian or Pacific Islander differed from AI on reservation, Hispanic, Non-Hispanic Asian and Non-Hispanic White.</p> <p>Non-Hispanic Asian differed from AI on reservation, Non-Hispanic Black and Native Hawaiian or Pacific Islander.</p> <p>Non-Hispanic White differed from AI on reservation, Non-Hispanic Black and Native Hawaiian or Pacific Islander.</p>
Other Residence Rate	<p>AI on reservation differed from Hispanic and Native Hawaiian or Pacific Islander.</p> <p>AI off reservation did not differ from the other categories.</p> <p>Hispanic differed from AI on reservation, Non-Hispanic Black and Non-Hispanic White.</p> <p>Non-Hispanic Black differed from Hispanic and Native Hawaiian and Pacific Islander.</p> <p>Native Hawaiian or Pacific Islander differed from AI on reservation, Non-Hispanic Black, Non-Hispanic Asian and Non-Hispanic White.</p> <p>Non-Hispanic Asian differed from Native Hawaiian or Pacific Islander.</p> <p>Non-Hispanic White differed from Hispanic and Native Hawaiian or Pacific Islander.</p>
Unresolved Rate	<p>AI on reservation differed from Hispanic.</p> <p>AI off reservation did not differ from other categories.</p> <p>Hispanic differed from AI on reservation, Non-Hispanic Asian and Non-Hispanic White.</p> <p>Non-Hispanic Black differed from Non-Hispanic White.</p> <p>Native Hawaiian or Pacific Islander did not differ from other categories.</p> <p>Non-Hispanic Asian differed from Hispanic and Non-Hispanic White.</p> <p>Non-Hispanic White differed from Hispanic, Non-Hispanic Black and Non-Hispanic Asian.</p>

Table A-1.D Percent of the E-Sample by Domain

Domain	Percent of E-Sample
1. AI On Reservation	0.2
2. AI Off Reservation	0.5
3. Hispanic	12.3
4. Non-Hispanic Black	11.8
5. Native Hawaiian or Pacific Islanders	0.2
6. Non-Hispanic Asian	3.6
7. Non-Hispanic White	71.4
Total	100.0

Age/Sex

Table A-2.A Types of Erroneous Enumerations by Age/Sex							
	Duplicate	Fictitious	Geocoding Error	Insufficient Information	Other Residence	Unresolved	Total
0-17	0.6	0.2	0.3	1.7	0.7	0.6	4.1
18-29 Male	1.0	0.5	0.2	2.0	2.1	1.4	7.1
18-29 Female	0.9	0.4	0.2	1.9	1.8	1.2	6.4
30-49 Male	0.8	0.3	0.2	1.9	0.8	0.7	4.8
30-49 Female	0.7	0.3	0.3	1.7	0.6	0.5	4.0
50+ Male	0.8	0.2	0.2	1.9	1.2	0.4	4.7
50+ Female	0.8	0.1	0.2	1.9	1.1	0.3	4.5
Total	0.8	0.3	0.2	1.8	1.0	0.6	4.7

Table A-2.B Standard Errors of Types of Erroneous Enumerations by Age/Sex

	Duplicate	Fictitious	Geocoding Error	Insufficient Information	Other Residence	Unresolved	Total
0-17	0.03	0.02	0.05	0.05	0.03	0.03	0.09
18-29 Male	0.06	0.05	0.04	0.10	0.08	0.06	0.16
18-29 Female	0.06	0.05	0.04	0.09	0.07	0.06	0.15
30-49 Male	0.04	0.03	0.05	0.06	0.03	0.03	0.11
30-49 Female	0.03	0.03	0.05	0.05	0.03	0.02	0.09
50+ Male	0.04	0.02	0.04	0.06	0.06	0.02	0.11
50+ Female	0.06	0.02	0.04	0.05	0.05	0.02	0.10
Total	0.03	0.02	0.04	0.03	0.02	0.02	0.07

Table A-2.C Significant Differences among Age/Sex

Rates	Which values are significant ($ t > 2.807$)
Total Erroneous Enumeration Rate	<p>0-17 differed from all other categories except 30-49 female.</p> <p>18-29 male differed from all other categories.</p> <p>18-29 female differed from all other categories.</p> <p>30-49 male differed from 0-17, 18-29 male, 18-29 female and 30-49 female.</p> <p>30-49 female differed from all other categories except 0-17.</p> <p>50+ male differed from 0-17, 18-29 male, 18-29 female and 30-49 female.</p> <p>50+ female differed from 0-17, 18-29 male, 18-29 female and 30-49 female.</p>
Duplicate Rate	<p>0-17 differed from all other categories.</p> <p>18-29 male differed from 0-17, 30-49 male, 30-49 female, and 50+ male.</p> <p>18-29 female differed from 0-17 and 30-49 females.</p> <p>30-49 male differed from 0-17, 18-29 male and 30-49 female.</p> <p>30-49 female differed from 0-17, 18-29 male, 18-29 female and 30-49 male.</p> <p>50+ male differed from 0-17 and 18-29 male.</p> <p>50+ female differed from 0-17.</p>

Fictitious Rate	<p>0-17 differed from all other categories except 30-49 female.</p> <p>18-29 male differed from all other categories except 18-29 female.</p> <p>18-29 female differed from 0-17, 30-49 female, 50+ male and 50+ female.</p> <p>30-49 male differed from 0-17, 18-29 male, 50+ male and 50+ female.</p> <p>30-49 female differed from 18-29 male, 18-29 female, 50+ male and 50+ female.</p> <p>50+ male differed from all other categories.</p> <p>50+ female differed from all other categories.</p>
Geocoding Rate	None
Insufficient Information Rate	<p>0-17 differed from 18-29 males.</p> <p>18-29 male differed from 0-17 and 30-49 female.</p> <p>18-29 female did not differ from any of the others.</p> <p>30-49 male differed from 30-49 female.</p> <p>30-49 female differed from 18-29 male, 30-49 male, 50+ male and 50+ female.</p> <p>50+ male differed from 30-49 female.</p> <p>50+ female differed from 30-49 female.</p>
Other Residence Rate	<p>0-17 differed from all other categories except 30-49 male.</p> <p>18-29 male differed from all other categories.</p> <p>18-29 female differed from all other categories.</p> <p>30-49 male differed from all other categories except 0-17.</p> <p>30-49 female differed from all other categories.</p> <p>50+ male differed from all other categories except 50+ female.</p> <p>50+ female differed from all other categories except 50+ male.</p>
Unresolved Rate	<p>0-17 differed from all other categories except 30-49 female.</p> <p>18-29 male differed from all other categories.</p> <p>18-29 female differed from all other categories.</p> <p>30-49 male differed from all other categories.</p> <p>30-49 female differed from all other categories except 0-17.</p> <p>50+ male differed from all other categories.</p> <p>50+ female differed from all other categories.</p>

Table A-2.D Percent of the E-Sample by Age/Sex

Age/Sex	Percent of E-Sample
0-17	25.7
18-29 Male	7.8
18-29 Female	7.8
30-49 Male	15.3
30-49 Female	15.8
50+ Male	12.6
50+ Female	15.0
Total	100.0

Tenure

Table A-3.A Types of Erroneous Enumerations by Tenure

	Duplicate	Fictitious	Geocoding Error	Insufficient Information	Other Residence	Unresolved	Total
Owner	0.6	0.1	0.3	1.4	1.0	0.3	3.6
Non-Owner	1.2	0.6	0.2	2.8	1.1	1.4	7.3
Total	0.8	0.3	0.2	1.8	1.0	0.6	4.7

Table A-3.B Standard Errors of Types of Erroneous Enumerations by Tenure

	Duplicate	Fictitious	Geocoding Error	Insufficient Information	Other Residence	Unresolved	Total
Owner	0.02	0.02	0.05	0.03	0.03	0.01	0.08
Non-Owner	0.06	0.04	0.04	0.08	0.04	0.06	0.13
Total	0.03	0.02	0.04	0.03	0.02	0.02	0.07

Table A-3.C Significant Differences among Tenure

Rates	Which values are significant ($t > 1.645$)
Total Erroneous Enumeration Rate	Significant (t=25.722)
Duplicate Rate	Significant (t=11.339)
Fictitious Rate	Significant (t=10.602)
Geocoding Rate	Not Significant (t=-0.834)
Insufficient Information Rate	Significant (t=17.488)
Other Residence Rate	Significant (t=1.658)
Unresolved Rate	Significant (t=18.101)

Table A-3.D Percent of the E-Sample by Tenure

Tenure	Percent of E-Sample
Owner	69.7
Non-Owner	30.3
Total	100.0

Return Rate

Table A-4.A Types of Erroneous Enumerations by Return Rate

	Duplicate	Fictitious	Geocoding Error	Insufficient Information	Other Residence	Unresolved	Total
High	0.6	0.2	0.2	1.7	1.0	0.5	4.2
Low	1.3	0.4	0.2	2.1	1.2	0.9	6.1
Not Applicable	1.0	0.4	0.3	2.0	1.1	0.8	5.5
Total	0.8	0.3	0.2	1.8	1.0	0.6	4.7

Table A-4.B Standard Errors of Types of Erroneous Enumerations by Return Rate

	Duplicate	Fictitious	Geocoding Error	Insufficient Information	Other Residence	Unresolved	Total
High	0.03	0.02	0.05	0.04	0.03	0.02	0.08
Low	0.06	0.04	0.05	0.07	0.05	0.06	0.14
Not Applicable	0.15	0.07	0.14	0.14	0.08	0.08	0.27
Total	0.03	0.02	0.04	0.03	0.02	0.02	0.07

Table A-4.C Significant Differences among Return Rate

Rates	Which values are significant ($t > 2.114$)
Total Erroneous Enumeration Rate	High return rate is lower than the other categories.
Duplicate Rate	High return rate is lower than the other categories.
Fictitious Rate	High return rate is lower than the other categories.
Geocoding Rate	None
Insufficient Information Rate	High return rate is lower than the Low return rate .
Other Residence Rate	High return rate is lower than the Low return rate .
Unresolved Rate	High return rate is lower than the other categories.

Table A.4.D Percent of the E-Sample by Return Rate

Return Rate	Percent of E-Sample
High	72.0
Low	23.5
Not Applicable	4.5
Total	100.0

MSA/TEA

Table A-5.A Types of Erroneous Enumerations by MSA/TEA

	Duplicate	Fictitious	Geocoding Error	Insufficient Info	Other Residence	Unresolved	Total
Large MSA MO/MB	0.9	0.4	0.3	2.1	0.9	0.6	5.2
Medium MSA MO/MB	0.5	0.2	0.2	1.9	1.0	0.7	4.5
Small MSA & Non-MSA MO/MB	0.6	0.2	0.3	1.6	1.0	0.6	4.4
All Other TEAs	1.0	0.1	0.1	1.4	1.5	0.5	4.6
Total	0.8	0.3	0.2	1.8	1.0	0.6	4.7

Table A-5.B Standard Errors of Types of Erroneous Enumerations by MSA/TEA

	Duplicate	Fictitious	Geocoding Error	Insufficient Info	Other Residence	Unresolved	Total
Large MSA MO/MB	0.05	0.04	0.11	0.07	0.04	0.03	0.16
Medium MSA MO/MB	0.04	0.03	0.06	0.06	0.04	0.05	0.12
Small MSA & Non-MSA MO/MB	0.05	0.06	0.08	0.07	0.05	0.03	0.15
All Other TEAs	0.06	0.02	0.02	0.07	0.07	0.04	0.13
Total	0.03	0.02	0.04	0.03	0.02	0.02	0.07

Table A-5.C Significant Differences among MSA/TEA

Rates	Which values are significant ($t > 2.378$)
Total Erroneous Enumeration Rate	Large MSA MO/MB is higher than all other categories.
Duplicate Rate	Large MSA MO/MB and All Other TEAs are higher than Medium MSA MO/MB and Small MSA & Non-MSA MO/MB
Fictitious Rate	Large MSA MO/MB is higher than all other categories. All Other TEAs differed from Medium MSA MO/MB .
Geocoding Rate	All Other TEAs were lower than all other categories.
Insufficient Information Rate	Large MSA MO/MB is higher than Small MSA & Non-MSA MO/MB and All Other TEAs . All Other TEAs is lower than Medium MSA MO/MB .
Other Residence Rate	All Other TEAs is higher than all other categories. Small MSA & Non-MSA MO/MB is higher than Large MSA MO/MB .
Unresolved Rate	All Other TEAs is lower than Large MSA MO/MB and Medium MSA MO/MB . Small MSA & Non-MSA MO/MB is lower than Medium MSA MO/MB .

Table A-5.D Percent of the E-Sample by MSA/TEA

MSA/TEA	Percent of E-Sample
Large MSA MO/MB	30.2
Medium MSA MO/MB	31.5
Small MSA & Non-MSA MO/MB	20.4
All Other TEAs	17.9
Total	100.0

Region

Table A-6.A Types of Erroneous Enumerations by Region

	Duplicate	Fictitious	Geocoding Error	Insufficient Information	Other Residence	Unresolved	Total
Northeast	1.3	0.3	0.3	1.8	1.0	0.4	5.0
Midwest	0.6	0.2	0.2	1.5	0.8	0.5	3.8
South	0.7	0.3	0.3	1.9	1.2	0.7	5.1
West	0.6	0.2	0.2	2.1	1.0	0.7	4.8
Total	0.8	0.3	0.2	1.8	1.0	0.6	4.7

Table A-6.B Standard Errors of Types of Erroneous Enumerations by Region

	Duplicate	Fictitious	Geocoding Error	Insufficient Information	Other Residence	Unresolved	Total
Northeast	0.08	0.05	0.07	0.08	0.05	0.03	0.16
Midwest	0.04	0.04	0.07	0.06	0.04	0.05	0.13
South	0.04	0.04	0.09	0.06	0.04	0.03	0.14
West	0.05	0.03	0.06	0.08	0.06	0.05	0.15
Total	0.03	0.02	0.04	0.03	0.02	0.02	0.07

Table A-6.C Significant Differences among Regions	
Rates	Which values are significant ($t > 2.378$)
Total Erroneous Enumeration Rate	Midwest is lower than all of the other regions.
Duplicate Rate	Northeast is higher than all of the other regions. West differs from South .
Fictitious Rate	None
Geocoding Rate	None
Insufficient Information Rate	Midwest is lower than all of the other regions. West differs Northeast
Other Residence Rate	Midwest is lower than Northeast and South . West differs from South .
Unresolved Rate	South and West differed from Northeast and Midwest .

Table A-6.D Percent of the E-Sample by Region	
Region	Percent of E-Sample
Northeast	19.1
Midwest	22.8
South	35.6
West	22.5
Total	100.0

Response Method

Table A-7.A Types of Erroneous Enumerations by Response Method

	Duplicate	Fictitious	Geocoding Error	Insuff. Info	Other Residence	Unresolved	Total
Self-reporting	0.4	0.1	0.2	0.9	0.9	0.4	2.9
Enumerator	1.9	0.9	0.3	4.7	1.5	1.2	10.6
Household Member	1.6	0.9	0.3	1.8	1.4	1.1	7.1
Proxy	3.9	1.1	0.2	27.4	2.6	19	37.2
Total	0.8	0.3	0.2	1.8	1.0	0.6	4.7

Table A-7.B Standard Errors of Types of Erroneous Enumerations by Response Method

	Duplicate	Fictitious	Geocoding Error	Insuff. Info	Other Residence	Unresolved	Total
Self-reporting	0.02	0.01	0.04	0.02	0.03	0.02	0.06
Enumerator	0.08	0.08	0.05	0.12	0.05	0.06	0.19
Household Member	0.07	0.08	0.06	0.06	0.05	0.06	0.16
Proxy	0.43	0.17	0.07	0.75	0.20	0.14	0.81
Total	0.03	0.02	0.04	0.03	0.02	0.02	0.07

Table A-7.C Significant Differences among Response Methods (Self-reporting, Enumerator returns completed with a household member and enumerator returns complete with a proxy respondent)

Rates	Which values are significant ($t > 2.114$)
Total Erroneous Enumeration Rate	All significantly different.
Duplicate Rate	All significantly different.
Fictitious Rate	Mail return is lower than both enumerator returns completed by a household member and enumerator returns completed by a proxy respondent
Geocoding Rate	None
Insufficient Information Rate	All significantly different.
Other Residence Rate	All significantly different.
Unresolved Rate	All significantly different.

Table A-7.D Percent of the E-Sample by Response Method

Response Method	Percent of E-Sample
Self-reporting	76.3
Enumerator	23.7
Household member	21.0
Proxy	2.7
Total	100.0

Imputation

Table A-8.A Types of Erroneous Enumerations by Imputation

	Duplicate	Fictitious	Geocoding Error	Insuff. Info	Other Residence	Unresolved	Total
No Imputation	0.6	0.2	0.3	0.6	1.0	0.6	3.3
Some Imputation	1.7	0.5	0.2	9.7	1.5	0.9	14.5
Total	0.8	0.3	0.2	1.8	1.0	0.6	4.7

Table A-8.B Standard Errors of Types of Erroneous Enumerations by Imputation

	Duplicate	Fictitious	Geocoding Error	Insuff. Info	Other Residence	Unresolved	Total
No Imputation	0.02	0.02	0.04	0.02	0.02	0.02	0.07
Some Imputation	0.10	0.04	0.05	0.21	0.07	0.04	0.25
Total	0.03	0.02	0.04	0.03	0.02	0.02	0.07

Table A-8.C Significant Differences among Imputation

Rates	Which values are significant (t >1.645)
Total Erroneous Enumeration Rate	Significant (t=44.489)
Duplicate Rate	Significant (t=10.781)
Fictitious Rate	Significant (t=5.276)
Geocoding Rate	Not Significant (t=0.461)
Insufficient Information Rate	Significant (t=42.948)
Other Residence Rate	Significant (t=8.480)
Unresolved Rate	Significant (t=8.824)

Table A-8.D Percent of the E-Sample by Imputation

	Percent of E-Sample
No Imputation	87.0
Some Imputation	13.0
Total	100.0

Housing Unit Match Status

Table A-9.A Types of Erroneous Enumerations by Housing Unit Match Status

	Duplicate	Fictitious	Geocoding Error	Insufficient Information	Other Residence	Unresolved	Total
Correctly Enumerated	0.5	0.3	0.0	1.7	1.0	0.6	4.0
Erroneous Enumerated	4.4	0.5	3.9	2.7	1.6	1.4	14.5
Unresolved Enumeration Status	4.8	0.5	0.0	4.5	2.5	0.6	12.9
No Matching	3.5	0.4	2.0	3.1	1.9	1.1	11.9
Total	0.8	0.3	0.2	1.8	1.0	0.6	4.7

Table A-9.B Standard Error of Types of Erroneous Enumerations by Housing Unit Match Status

	Duplicate	Fictitious	Geocoding Error	Insufficient Information	Other Residence	Unresolved	Total
Correctly Enumerated	0.02	0.02	0.00	0.03	0.02	0.02	0.06
Erroneous Enumerated	0.37	0.10	0.60	0.29	0.14	0.14	0.75
Unresolved Enumeration Status	1.58	0.34	0.00	1.35	1.19	0.28	2.71
No Matching	0.31	0.10	1.05	0.32	0.20	0.12	1.05
Total	0.03	0.02)	0.04	0.03	0.02	0.02	0.07

Table A-9.C Significant Differences among Housing Unit Match Status

Rates	Which values are significant ($t > 2.378$)
Total Erroneous Enumeration Rate	Correctly enumerated housing units were lower than all other categories.
Duplicate Rate	Correctly enumerated housing units were lower than all other categories.
Fictitious Rate	Correctly enumerated housing units were lower erroneously enumerated housing units.
Geocoding Rate	Erroneously enumerated housing units were higher than correctly enumerated housing units and housing units with unresolved enumeration status .
Insufficient Information Rate	Correctly enumerated housing units were lower than erroneously enumerated housing units and housing units with no matching .
Other Residence Rate	Correctly enumerated housing units were lower than erroneously enumerated housing units and housing units with no matching .
Unresolved Rate	Correctly enumerated housing units were lower than erroneously enumerated housing units and housing units with no matching . Housing units with unresolved enumeration status were lower than erroneously enumerated housing units.

Table A-9.D Percent of the E-Sample by Housing Unit Match Status

Housing Unit Match Status	Percent of E-Sample (People)
Correctly Enumerated	92.4
Erroneous Enumerated	4.6
Unresolved Enumeration Status	0.1
No Matching	2.9
Total	100.0

Initial Address Status

Table A-10.A Type of Erroneous Enumeration by Initial Address Status							
	Duplicate	Fictitious	Geocoding Error	Insuff. Info	Other Residence	Unresolved	Total
HU matched during HU phase	0.4	0.1	0.0	1.6	0.9	0.4	3.4
HU not matched during HU phase	3.9	0.5	2.8	2.8	1.7	1.3	13.0
HU added to DMAF after HU phase of ACE	3.5	0.4	2.0	3.1	1.9	1.1	11.9
Conflicting households	5.7	8.8	0.0	3.7	7.1	11.2	36.5
Total	0.8	0.3	0.2	1.8	1.0	0.6	4.7

Table A-10.B Standard Errors of Type of Erroneous Enumeration by Initial Address Status

	Duplicate	Fictitious	Geocoding Error	Insuff. Info	Other Residence	Unresolved	Total
HU matched during HU phase	0.01	0.01	0.00	0.03	0.02	0.02	0.05
HU not matched during HU phase	0.28	0.08	0.43	0.22	0.13	0.11	0.57
HU added to DMAF after HU phase of ACE	0.31	0.10	1.05	0.32	0.20	0.12	1.05
Conflicting households	0.43	0.89	0.00	0.31	0.48	0.66	1.09
Total	0.03	0.02	0.04	0.03	0.02	0.02	0.07

Table A-10.C Significant Differences among Initial Address Status

Rates	Which values are significant ($ t > 2.378$)
Total Erroneous Enumeration Rate	<p>Housing unit matched is lower than housing unit not matched, housing unit added and conflicting households.</p> <p>Housing unit not matched is higher than housing unit matched and lower than conflicting household.</p> <p>Housing unit added is higher than housing unit matched and lower than conflicting household.</p> <p>Conflicting household is higher than all other categories.</p>
Duplicate Rate	<p>Housing unit matched is lower than housing unit not matched, housing unit added and conflicting households.</p> <p>Housing unit not matched is higher than housing unit matched and lower than conflicting household.</p> <p>Housing unit added is higher than housing unit matched and lower than conflicting household.</p> <p>Conflicting household is higher than all other categories.</p>

Fictitious Rate	<p>Housing unit matched is lower than housing unit not matched, housing unit added and conflicting households.</p> <p>Housing unit not matched is higher than housing unit matched and lower than conflicting household.</p> <p>Housing unit added is higher than housing unit matched and lower than conflicting household.</p> <p>Conflicting household is higher than all other categories.</p>
Geocoding Rate	<p>Housing unit not matched is higher than housing unit matched and conflicting household.</p>
Insufficient Information Rate	<p>Housing unit matched is lower than all the other categories.</p>
Other Residence Rate	<p>Housing unit matched is lower than housing unit not matched, housing unit added and conflicting households.</p> <p>Housing unit not matched is higher than housing unit matched and lower than conflicting household.</p> <p>Housing unit added is higher than housing unit matched and lower than conflicting household.</p> <p>Conflicting household is higher than all other categories.</p>
Unresolved Rate	<p>Housing unit matched is lower than housing unit not matched, housing unit added and conflicting households.</p> <p>Housing unit not matched is higher than housing unit matched and lower than conflicting household.</p> <p>Housing unit added is higher than housing unit matched and lower than conflicting household.</p> <p>Conflicting household is higher than all other categories.</p>

Table A-10.D Percent of the E-Sample by initial Address Status

Final Address Status	Percent of E-Sample
HU matched	89.3
HU not matched	6.4
HU added to DMAF after HU phase of ACE	2.9
Conflicting households	1.4
Total	100.0

Number of units at the basic street address

Table A-11.A Types of Erroneous Enumerations by Number of Units at the Basic Street Address

	Duplicate	Fictitious	Geocoding Error	Insufficient Information	Other Residence	Unresolved	Total
1 unit	0.4	0.2	0.2	1.5	1.0	0.4	3.7
2-9 units	3.3	0.6	0.1	3.0	1.2	1.3	9.5
10+ units	1.0	0.5	0.3	3.0	1.1	1.6	7.6
Total	0.8	0.3	0.2	1.8	1.0	0.6	4.7

Table A-11.B Standard Error of Types of Erroneous Enumerations by Number of Units at the Basic Street Address

	Duplicate	Fictitious	Geocoding Error	Insufficient Information	Other Residence	Unresolved	Total
1 unit	0.02	0.02	0.05	0.03	0.03	0.01	0.07
2-9 units	0.15	0.08	0.05	0.12	0.08	0.10	0.24
10+ units	0.11	0.06	0.10	0.16	0.08	0.11	0.26
Total	0.03	0.02	0.04	0.03	0.02	0.02	0.07

Table A-11.C Significant Differences among number of units at the basic street address

Rates	Which values are significant ($t > 2.114$)
Total Erroneous Enumeration Rate	All significantly different.
Duplicate Rate	All significantly different.
Fictitious Rate	1 unit has a lower rate than 2-9 units and 10+ units .
Geocoding Rate	None
Insufficient Information Rate	1 unit has a lower rate than 2-9 units and 10+ units .
Other Residence Rate	1 unit has a lower rate than 2-9 units .
Unresolved Rate	All significantly different.

Table A-11.D Percent of the E-Sample by Units at Basic Street Address

UBSA	Percent of E-Sample
1 unit	78.3
2-9 units	10.1
10+ units	11.6
Total	100.0

Whole/Partial Match Code

Table A-12.A Types of Erroneous Enumerations by Whole/Partial Match Code

	Duplicate	Fictitious	Geocoding Error	Insufficient Information	Other Residence	Unresolved	Total
Partial household match	1.20	0.34	0.00	0.51	4.19	1.81	8.04
Whole household nonmatch	4.92	1.85	1.91	1.60	4.71	3.46	18.45
Whole household match	0.03	0.00	0.00	0.58	0.03	0.00	0.64
Other Residence Other	0.00	0.00	0.00	99.76	0.00	0.04	99.8
Total	0.8	0.3	0.2	1.8	1.0 (.02)	0.6	4.7

Table A-12.B Standard Errors of Types of Erroneous Enumerations by Whole/Partial Match Code

	Duplicate	Fictitious	Geocoding Error	Insufficient Information	Other Residence	Unresolved	Total
Partial household match	0.07	0.03	0.00	0.04	0.10	0.08	0.15
Whole household nonmatch	0.20	0.16	0.33	0.07	0.17	0.14	0.47
Whole household match	0.00	0.00	0.00	0.02	0.00	0.00	0.02
Other Residence Other	0.00	0.00	0.00	0.06	0.00	0.02	0.06
Total	0.03	0.02	0.04	0.03	0.02	0.02	0.07

Table A-12.C Percent of the E-Sample by Whole/Partial Match Code

Response Method	Percent of E-Sample
Partial household match	10.0
Whole household nonmatch	12.6
Whole household match	76.3
All others	1.1
Total	100.0

Form length

Table A-13.A Types of Erroneous Enumerations by Form Length

	Duplicate	Fictitious	Geocoding Error	Insufficient Information	Other Residence	Unresolved	Total
Short	0.7	0.3	0.2	1.8	1.1	0.6	4.7
Long	0.8	0.3	0.2	1.8	0.9	0.5	4.6
Total	0.8	0.3	0.2	1.8	1.0	0.6	4.7

Table A-13.B Standard Errors of Types of Erroneous Enumerations by Form Length

	Duplicate	Fictitious	Geocoding Error	Insufficient Information	Other Residence	Unresolved	Total
Short	0.03	0.02	0.04	0.04	0.03	0.02	0.08
Long	0.05	0.03	0.04	0.07	0.04	0.03	0.11
Total	0.03	0.02	0.04	0.03	0.02	0.02	0.07

Table A-13.C Significant Differences among Form Length

Rates	Which values are significant ($t > 1.645$)
Total Erroneous Enumeration Rate	Not significant ($t=1.568$)
Duplicate Rate	Significant ($t=1.696$)
Fictitious Rate	Not significant ($t=0.186$)
Geocoding Rate	Not significant ($t=0.121$)
Insufficient Information Rate	Not significant ($t=0.667$)
Other Residence Rate	Significant ($t=3.597$)
Unresolved Rate	Significant ($t=4.116$)

Table A-13.D Percent of the E-Sample by Form Length

Form Length	Percent of E-Sample
Short	83.5
Long	16.5
Total	100.0

Non-response follow up

Table A-14.A Types of Erroneous Enumerations by Non-Response Follow Up

	Duplicate	Fictitious	Geocoding Error	Insufficient Information	Other Residence	Unresolved	Total
Not in NRFU universe	0.5	0.1	0.2	1.0	0.9	0.4	3.0
In NRFU universe	1.5	0.8	0.3	4.0	1.5	1.1	9.2
Total	0.8	0.3	0.2	1.8	1.0	0.6	4.7

Table A-14.B Standard Errors of Types of Erroneous Enumerations by Non-Response Follow Up

	Duplicate	Fictitious	Geocoding Error	Insufficient Information	Other Residence	Unresolved	Total
Not in NRFU universe	0.02	0.01	0.04	0.03	0.03	0.02	0.06
In NRFU universe	0.07	0.07	0.05	0.10	0.05	0.05	0.16
Total	0.03	0.02	0.04	0.03	0.02	0.02	0.07

Table A-14.C Significant Differences among Non-Response Follow Up

Rates	Which values are significant ($t > 1.645$)
Total Erroneous Enumeration Rate	Significant (t=38.220)
Duplicate Rate	Significant (t=16.001)
Fictitious Rate	Significant (t=10.124)
Geocoding Rate	Not significant (t=1.176)
Insufficient Information Rate	Significant (t=30.647)
Other Residence Rate	Significant (t=11.968)
Unresolved Rate	Significant (t=13.437)

Table A-14.D Percent of the E-Sample by Non-Response Follow Up

	Percent of E-Sample
Not in NRFU universe	71.9
In NRFU universe	28.1
Total	100.0

Coverage Edit Follow Up

Table A.15.A Type of Erroneous Enumeration by Coverage Edit Follow Up

	Duplicate	Fictitious	Geocoding Error	Insuff. Info	Other Residence	Unresolved	Total
No Coverage Edit Follow Up	0.7	0.3	0.2	1.7	1.0	0.6	4.6
Coverage Edit Follow Up	1.3	0.1	0.2	3.2	1.5	0.8	7.1
Total	0.8	0.3	0.2	1.8	1.0	0.6	4.7

Table A.15.B Standard Errors of Type of Erroneous Enumeration by Coverage Edit Follow Up

	Duplicate	Fictitious	Geocoding Error	Insuff. Info	Other Residence	Unresolved	Total
No Coverage Edit Follow Up	0.03	0.02	0.04	0.03	0.02	0.02	0.07
Coverage Edit Follow Up	0.12	0.02	0.08	0.18	0.12	0.09	0.29
Total	0.03	0.02	0.04	0.03	0.02	0.02	0.07

Table A.15.C Significant Differences among Coverage Edit Follow Up

Rates	Which values are significant ($t > 1.645$)
Total Erroneous Enumeration Rate	Significant (t=8.727)
Duplicate Rate	Significant (t=4.848)
Fictitious Rate	Significant (t=6.750)
Geocoding Rate	Not significant (t=1.194)
Insufficient Information Rate	Significant (t=7.898)
Other Residence Rate	Significant (t=4.326)
Unresolved Rate	Significant (t=2.179)

Table A.15.D Percent of the E-Sample by Coverage Edit Follow Up

Coverage Edit Follow Up	Percent of E-Sample
No Coverage Edit Follow Up	95.5
Coverage Edit Follow Up	4.5
Total	100.0

Coverage Improvement Follow Up

Table A-16.A Types of Erroneous Enumerations by Coverage Improvement Follow Up

	Duplicate	Fictitious	Geocoding Error	Insufficient Information	Other Residence	Unresolved	Total
No Coverage Improvement Follow Up	0.7	0.2	0.2	1.6	1.0	0.6	4.3
Coverage Improvement Follow Up	6.1	1.2	0.3	13.5	3.0	1.8	25.9
Total	0.8	0.3	0.2	1.8	1.0	0.6	4.7

Table A-16.B Standard Errors of Types of Erroneous Enumerations by Coverage Improvement Follow Up

	Duplicate	Fictitious	Geocoding Error	Insufficient Information	Other Residence	Unresolved	Total
No Coverage Improvement Follow Up	0.02	0.02	0.04	0.03	0.02	0.02	0.07
Coverage Improvement Follow Up	0.49	0.21	0.13	0.65	0.27	0.21	0.85
Total	0.03	0.02	0.04	0.03	0.02	0.02	0.07

Table A-16.C Significant Differences among Coverage Improvement Follow Up

Rates	Which values are significant ($t > 1.645$)
Total Erroneous Enumeration Rate	Significant (t=25.390)
Duplicate Rate	Significant (t=11.111)
Fictitious Rate	Significant (t=4.528)
Geocoding Rate	Not significant (t=0.643)
Insufficient Information Rate	Significant (t=18.332)
Other Residence Rate	Significant (t=7.420)
Unresolved Rate	Significant (t=5.618)

**Table A-16.D Percent of the E-Sample by
Coverage Improvement Follow Up**

Coverage Improvement Follow Up	Percent of E-Sample
No Coverage Improvement Follow Up	98.2
Coverage Improvement Follow Up	1.8
Total	100.0

Capture Method

Table A-17.A Types of Erroneous Enumerations by Capture Method

	Duplicate	Fictitious	Geocoding Error	Insufficient Information	Other Residence	Unresolved	Total
Scanned	0.7	0.3	0.2	1.8	1.0	0.6	4.6
Not Scanned	1.7	0.8	0.2	5.0	1.4	0.9	10.0
Total	0.8	0.3	0.2	1.8	1.0	0.6	4.7

Table A-17.B Standard Errors of Types of Erroneous Enumerations by Capture Method

	Duplicate	Fictitious	Geocoding Error	Insufficient Information	Other Residence	Unresolved	Total
Scanned	0.02	0.02	0.04	0.03	0.02	0.02	0.07
Not Scanned	0.23	0.19	0.10	0.38	0.19	0.13	0.53
Total	0.03	0.02	0.04	0.03	0.02	0.02	0.07

Table A-17.C Significant Differences among Capture Method

Rates	Which values are significant ($t > 1.645$)
Total Erroneous Enumeration Rate	Significant (t=10.099)
Duplicate Rate	Significant (t=3.866)
Fictitious Rate	Significant (t=2.972)
Geocoding Rate	Not significant (t=0.127)
Insufficient Information Rate	Significant (t=8.5000)
Other Residence Rate	Significant (t=1946)
Unresolved Rate	Significant (t=1.912)

Table A-17.D Percent of the E-Sample by Capture System

Capture System	Percent of E-Sample
Scanned	98.5
Not Scanned	1.5
Total	100.0

Table A-18.A Types of Erroneous Enumerations by Data Capture Center

	Duplicate	Fictitious	Geocoding	Insufficient	Other	Unresolved	Total
			Error	Information	Residence		
Missing	1.54	0.00	0.00	0.67	1.01	0.76	3.97
Baltimore	0.85	0.26	0.23	1.47	0.91	0.47	4.19
J-ville	0.93	0.34	0.30	2.00	0.89	0.61	5.06
Phoenix	0.81	0.31	0.31	1.87	1.23	0.71	5.23
Pomona	0.56	0.20	0.16	1.99	0.99	0.68	4.59
Total	0.8	0.3	0.2	1.8	1.0	0.6	4.7

Table A-18.B Standard Errors of Types of Erroneous Enumerations by Data Capture Center

	Duplicate	Fictitious	Geocoding	Insufficient	Other	Unresolved	Total
			Error	Information	Residence		
Missing	0.63	0.00	0.00	0.36	0.45	0.41	0.92
Baltimore	0.06	0.03	0.06	0.05	0.04	0.04	0.12
J-ville	0.08	0.08	0.11	0.11	0.06	0.07	0.21
Phoenix	0.04	0.05	0.11	0.06	0.05	0.04	0.15
Pomona	0.04	0.03	0.04	0.07	0.04	0.04	0.12
Total	0.03	0.02	0.04	0.03	0.02	0.02	0.07

Table A-18.C Percent of the E-Sample by Data Capture Center

Data Capture Center	Percent of E-Sample
Missing	0.1
Baltimore	29.0
Jeffersonville	9.5
Phoenix	31.1
Pomona	30.3
Total	100.0

A.C.E. person follow-up

Table A-19.A Type of Erroneous Enumerations by Follow-Up

	Duplicate	Fictitious	Geocoding Error	Insuf. Info	Other Residence	Unresolved	Total
No follow-up	0.8	0.0	0.2	2.2	0.0	0.0	3.2
Follow-up	0.6	1.6	0.4	0.0	6.1	3.6	12.3
Total	0.8	0.3	0.2	1.8	1.0	0.6	4.7

Table A-19.B Standard Errors of Type of Erroneous Enumerations by Follow-Up

	Duplicate	Fictitious	Geocoding Error	Insuf. Info	Other Residence	Unresolved	Total
No follow-up	0.03	0.00	0.03	0.04	0.00	0.00	0.06
Follow-up	0.05	0.12	0.19	0.01	0.15	0.12	0.30
Total	0.03	0.02	0.04	0.03	0.02	0.02	0.07

Table A-19.C Percent of the E-Sample by Person Follow-Up

Follow-up	Percent of E-Sample
No Follow-up	83.2
Follow-up	16.8
Total	100.0

A.C.E. relisted cluster

Table A-20.A Type of Erroneous Enumerations by Relist

	Duplicate	Fictitious	Geocoding Error	Insufficient Information	Other Residence	Unresolved	Total
Not relisted	0.8	0.3	0.2	1.8	1.0	0.6	4.7
Relisted	1.4	1.0	0.6	1.8	1.2	1.2	7.3
Total	0.8	0.3	0.2	1.8	1.0	0.6	4.7

Table A-20.B Standard Errors of Type of Erroneous Enumerations by Relist

	Duplicate	Fictitious	Geocoding Error	Insufficient Information	Other Residence	Unresolved	Total
Not relisted	0.03	0.02	0.04	0.03	0.02	0.02	0.07
Relisted	0.64	0.82	0.48	0.33	0.35	0.36	1.28
Total	0.03	0.02	0.04	0.03	0.02	0.02	0.07

Table A-20.C Percent of the E-Sample by Relist

Relist	Percent of E-Sample
Not Relisted	99.8
Relisted	0.2
Total	100.0

Table A-21.A Types of Erroneous Enumerations by ACE Regional Office

	Duplicate	Fictitious	Geocoding Error	Insufficient Information	Other Residence	Unresolved	Total
Boston	1.1	0.2	0.3	1.4	1.2	0.2	4.2
New York	2.0	0.6	0.1	2.3	1.0	0.7	6.7
Philly	0.6	0.3	0.4	1.8	0.9	0.5	4.5
Detroit	0.4	0.2	0.2	1.5	0.8	0.6	3.8
Chicago	0.7	0.3	0.3	1.6	0.8	0.5	4.2
Kansas City	0.6	0.1	0.1	1.3	1.0	0.6	3.8
Seattle	0.8	0.2	0.1	2.5	0.9	0.6	5.0
Charlotte	0.7	0.3	0.1	1.6	1.1	0.7	4.5
Atlanta	0.7	0.2	0.3	2.1	1.5	0.5	5.4
Dallas	0.8	0.5	0.5	1.8	1.0	0.8	5.5
Denver	0.5	0.3	0.3	1.8	1.3	0.8	5.0
Los Angeles	0.5	0.2	0.3	1.8	0.7	0.7	4.2
Total	0.8	0.3	0.2	1.8	1.0	0.6	4.7

Table A-21.B Standard Errors of Types of Erroneous Enumerations by ACE Regional Office

	Duplicate	Fictitious	Geocoding Error	Insufficient Information	Other Residence	Unresolved	Total
Boston	0.16	0.05	0.10	0.12	0.09	0.03	0.25
New York	0.17	0.09	0.07	0.15	0.10	0.08	0.30
Philly	0.06	0.08	0.14	0.13	0.06	0.05	0.24
Detroit	0.06	0.04	0.14	0.09	0.06	0.11	0.22
Chicago	0.07	0.09	0.11	0.09	0.06	0.07	0.21
Kansas City	0.07	0.03	0.04	0.10	0.08	0.06	0.20
Seattle	0.11	0.04	0.05	0.18	0.07	0.06	0.25
Charlotte	0.07	0.11	0.04	0.09	0.08	0.06	0.20
Atlanta	0.07	0.04	0.13	0.12	0.10	0.05	0.24
Dallas	0.08	0.09	0.31	0.10	0.07	0.07	0.36
Denver	0.06	0.07	0.13	0.12	0.14	0.13	0.28
Los Angeles	0.05	0.05	0.12	0.11	0.06	0.07	0.20
Total	0.03	0.02	0.04	0.03	0.02	0.02	0.07

Race and Tenure

Table A-22.A Types of Erroneous Enumerations by Domain and Tenure							
	Duplicate	Fictitious	Geocoding Error	Insufficient Information	Other Residence	Imputed	Total
American Indian on Reservation, owner	1.0	0.1	0.2	0.8	1.8	0.6	4.4
American Indian on Reservation, non-owner	0.3	0.1	0.0	1.0	1.6	0.9	3.9
American Indian off Reservation, owner	0.6	0.2	0.2	2.9	1.1	0.5	5.5
American Indian off Reservation, non-owner	0.8	0.4	0.1	2.5	1.4	1.6	6.8
Hispanic, owner	0.8	0.3	0.2	1.3	0.8	0.5	3.8
Hispanic, non- owner	1.5	0.7	0.3	2.0	0.9	1.8	7.2
Black, owner	1.1	0.3	0.4	2.3	1.2	0.5	5.8
Black, non- owner	1.3	1.0	0.2	3.7	1.0	1.7	8.8
Pacific Islander, owner	1.0	0.3	0.0	3.7	0.8	0.4	6.2
Pacific Islander, non-owner	0.6	0.6	0.0	5.1	0.1	1.2	7.7
Asian, owner	0.8	0.2	0.5	1.3	1.1	0.3	4.2
Asian, non- owner	1.4	0.7	0.1	2.6	1.2	1.5	7.6
White, owner	0.5	0.1	0.2	1.3	1.0	0.3	3.3
White, non- owner	1.1	0.4	0.2	2.9	1.2	1.1	6.8
Total	0.8	0.3	0.2	1.8	1.0	0.6	4.7

Table A-22.B Standard Errors of Types of Erroneous Enumerations by Domain and Tenure							
	Duplicate	Fictitious	Geocoding Error	Insufficient Information	Other Residence	Imputed	Total
American Indian on Reservation, owner	0.20	0.04	0.13	0.21	0.34	0.13	0.43
American Indian on Reservation, non-owner	0.12	0.04	0.00	0.31	0.27	0.22	0.48
American Indian off Reservation, owner	0.20	0.11	0.14	0.64	0.27	0.22	0.74
American Indian off Reservation, non-owner	0.26	0.20	0.06	0.63	0.42	0.41	0.89
Hispanic, owner	0.08	0.07	0.07	0.10	0.08	0.06	0.19
Hispanic, non- owner	0.11	0.08	0.12	0.13	0.07	0.13	0.27
Black, owner	0.10	0.07	0.21	0.14	0.08	0.04	0.28
Black, non- owner	0.11	0.13	0.05	0.19	0.07	0.14	0.28
Pacific Islander, owner	0.32	0.20	0.00	0.88	0.37	0.23	1.07
Pacific Islander, non-owner	0.27	0.36	0.01	1.51	0.09	0.35	1.54
Asian, owner	0.16	0.06	0.26	0.15	0.12	0.06	0.37
Asian, non- owner	0.31	0.18	0.07	0.34	0.18	0.19	0.54
White, owner	0.02	0.01	0.04	0.03	0.03	0.01	0.07
White, non- owner	0.08	0.04	0.04	0.10	0.05	0.05	0.16
Total	0.03	0.02	0.04	0.03	0.02	0.02	0.07

Table A-23: Type of Erroneous Enumeration by Post-strata

Simulated Race Domain	Tenure	TEA	Mail Return Rate	Region	Dup.	Fict.	Geo. Error	Insuff. Info	Other Res.	Imput.	Total
White and Other	Owner	Large MSA MO/MB	High	Northeast	0.38	0.10	0.08	0.95	0.96	0.12	2.59
				Midwest	0.17	0.21	0.00	1.34	0.67	0.17	2.57
				South	0.15	0.06	0.96	1.20	0.66	0.30	3.34
				West	0.35	0.05	0.27	2.09	0.82	0.25	3.83
			Low	Northeast	1.31	0.76	0.93	1.80	1.00	0.31	6.11
				Midwest	0.95	0.06	0.00	1.65	0.88	0.60	4.14
				South	0.36	0.22	0.00	2.04	0.75	0.28	3.65
				West	0.73	0.00	0.60	1.75	1.02	0.49	4.60
		Medium MSA MO/MB	High	Northeast	0.30	0.00	0.39	0.82	0.83	0.16	2.49
				Midwest	0.15	0.05	0.26	0.90	0.56	0.17	2.08
				South	0.21	0.07	0.19	1.52	1.14	0.21	3.35
				West	0.30	0.06	0.03	1.38	1.02	0.23	3.02
			Low	Northeast	1.25	0.13	0.19	0.78	1.50	0.19	4.04
				Midwest	0.90	0.13	0.00	2.00	0.61	0.16	3.80
				South	0.79	0.24	0.06	1.69	1.12	0.48	4.37
				West	0.42	0.25	0.86	1.85	1.09	0.51	4.99

Table A-23: Type of Erroneous Enumeration by Post-strata

Simulated Race Domain	Tenure	TEA	Mail Return Rate	Region	Dup.	Fict.	Geo. Error	Insuff. Info	Other Res.	Imput.	Total
		Small MSA & Non-MSA MO/MB	High	Northeast	0.34	0.02	1.14	0.69	1.03	0.11	3.33
				Midwest	0.35	0.02	0.44	0.97	0.78	0.20	2.77
				South	0.35	0.21	0.15	1.13	0.99	0.24	3.07
				West	0.47	0.00	0.11	1.09	0.71	0.26	2.65
			Low	Northeast	0.61	0.00	0.00	0.75	0.95	0.17	2.48
				Midwest	1.41	0.04	0.41	1.50	0.71	0.41	4.48
				South	0.46	0.08	0.69	1.34	1.48	0.31	4.35
				West	0.41	0.00	0.21	1.97	1.27	0.36	4.23
		Other TEA	High	Northeast	0.50	0.04	0.08	0.96	1.16	0.08	2.83
				Midwest	0.51	0.04	0.00	0.72	1.16	0.31	2.73
				South	0.58	0.10	0.00	1.09	1.44	0.15	3.36
				West	0.38	0.03	0.29	1.99	1.76	0.40	4.85
			Low	Northeast	1.45	0.02	0.00	1.22	1.33	0.11	4.12
				Midwest	1.00	0.00	0.05	0.88	1.39	0.47	3.79
				South	1.10	0.10	0.06	1.36	1.40	0.41	4.44
				West	0.87	0.18	0.07	1.86	2.27	0.61	5.85

Table A-23: Type of Erroneous Enumeration by Post-strata

Simulated Race Domain	Tenure	TEA	Mail Return Rate	Region	Dup.	Fict.	Geo. Error	Insuff. Info	Other Res.	Imput.	Total
Black	Owner	Large MSA MO/MB	High	Region 1	0.84	0.50	0.28	3.04	0.85	0.83	6.35
			Low	Region 1	2.06	0.81	0.04	3.43	0.81	1.20	8.35
		Medium MSA MO/MB	High	Region 2	1.20	0.27	0.30	2.99	1.14	1.16	7.06
			Low	Region 2	0.96	0.42	0.21	3.46	1.18	2.06	8.29
		Small MSA & Non-MSA MO/MB	High	Region 3	0.82	0.26	0.09	2.55	1.00	0.90	5.61
			Low	Region 3	1.27	0.53	0.18	3.20	1.27	1.60	8.05
		Other TEA	High	Region 4	1.13	0.13	0.09	2.02	1.92	0.76	6.06
			Low	Region 4	1.68	0.45	0.01	2.73	1.74	1.31	7.93
	Renter	Large & Medium MSA MO/MB	High	Region 1	0.42	0.29	0.36	2.30	0.87	0.47	4.72
			Low	Region 1	2.95	0.36	0.02	3.44	1.29	0.71	8.78
		Small MSA & Non-MSA MO/MB & other TEA	High	Region 2	0.96	0.50	0.66	1.83	1.61	0.37	5.93
			Low	Region 2	1.87	0.13	0.09	1.82	1.42	0.51	5.83
		Large & Medium MSA MO/MB	High	Region 3	0.88	0.96	0.26	3.79	0.98	1.59	8.47
			Low	Region 3	1.96	1.11	0.03	4.08	0.91	2.15	10.24

Table A-23: Type of Erroneous Enumeration by Post-strata

Simulated Race Domain	Tenure	TEA	Mail Return Rate	Region	Dup.	Fict.	Geo. Error	Insuff. Info	Other Res.	Imput.	Total	
Hispanic		Small MSA & Non-MSA MO/MB & other TEA	High		1.55	1.00	0.07	3.04	1.12	1.51	8.29	
			Low		1.39	0.66	0.25	2.16	2.29	0.99	7.73	
	Owner	Large & Medium MSA MO/MB	High		0.48	0.11	0.17	1.27	0.71	0.43	3.16	
			Low		1.35	1.00	0.33	1.45	0.97	0.67	5.77	
		Small MSA & Non-MSA MO/MB & other TEA	High		0.60	0.15	0.15	1.23	0.98	0.42	3.54	
			Low		1.77	0.03	0.17	1.09	1.04	0.48	4.56	
		Renter	Large & Medium MSA MO/MB		High	1.09	0.47	0.39	1.96	0.73	1.57	6.22
					Low	2.13	1.27	0.38	2.03	1.20	2.21	9.21
			Small MSA & Non-MSA MO/MB & other TEA		High	1.70	0.56	0.10	1.99	1.34	2.01	7.70
					Low	2.35	0.79	0.09	2.65	0.48	1.51	7.87
Pacific Islander	Owner				0.95	0.34	0.00	3.66	0.82	0.44	6.21	
	Renter				0.58	0.64	0.01	5.10	0.14	1.20	7.67	

Table A-23: Type of Erroneous Enumeration by Post-strata

Simulated Race Domain	Tenure	TEA	Mail Return Rate	Region	Dup.	Fict.	Geo. Error	Insuff. Info	Other Res.	Imput.	Total
Asian	Owner				0.81	0.21	0.46	1.31	1.05	0.33	4.16
	Renter				1.38	0.74	0.13	2.61	1.22	1.47	7.55
American Indians on Reservations	Owner				0.95	0.07	0.15	0.78	1.78	0.62	4.35
	Renter				0.29	0.13	0.00	0.96	1.59	0.89	3.85
American Indians Not on Reservations	Owner				0.55	0.15	0.18	2.93	1.10	0.53	5.45
	Renter				0.80	0.44	0.09	2.49	1.39	1.63	6.84

Table A-24: Percentage of Types of Erroneous Enumerations by Domain and Region

Domain	Region	Duplicate	Fictitious	Geocoding Error	Insufficient Information	Other Residence	Unresolved	Total EE
American Indian off reservation	NE	1.91%	1.14%	0.00%	3.67%	0.35%	0.25%	7.31%
American Indian off reservation	MW	0.75%	0.00%	0.21%	1.29%	0.89%	1.58%	4.71%
American Indian off reservation	S	0.70%	0.30%	0.16%	2.38%	1.43%	1.03%	5.98%
American Indian off reservation	W	0.27%	0.20%	0.12%	3.86%	1.29%	0.80%	6.53%
American Indian on reservation	NE	0.00%	1.00%	0.00%	3.00%	0.00%	0.04%	4.04%
American Indian on reservation	MW	0.64%	0.04%	0.00%	0.64%	1.17%	1.00%	3.50%
American Indian on reservation	S	2.16%	0.11%	0.00%	0.32%	0.84%	0.73%	4.16%
American Indian on reservation	W	0.68%	0.08%	0.14%	0.89%	1.92%	0.65%	4.36%
Asian	NE	2.07%	0.62%	0.04%	1.96%	1.09%	0.86%	6.65%
Asian	MW	0.71%	0.46%	0.00%	1.92%	1.36%	0.93%	5.38%
Asian	S	0.61%	0.41%	1.17%	1.28%	1.25%	0.89%	5.62%
Asian	W	0.78%	0.30%	0.20%	1.91%	1.01%	0.61%	4.81%
Black	NE	2.36%	0.72%	0.04%	4.16%	1.16%	0.84%	9.29%
Black	MW	1.05%	0.99%	0.05%	2.71%	0.91%	1.51%	7.23%
Black	S	0.97%	0.55%	0.40%	2.54%	1.16%	0.96%	6.57%
Black	W	0.44%	0.52%	0.28%	3.89%	0.99%	1.45%	7.56%

Table A-24: Percentage of Types of Erroneous Enumerations by Domain and Region

Domain	Region	Duplicate	Fictitious	Geocoding Error	Insufficient Information	Other Residence	Unresolved	Total EE
Hispanic	NE	2.18%	0.63%	0.21%	1.77%	1.04%	1.13%	6.96%
Hispanic	MW	1.71%	0.53%	0.11%	1.70%	1.06%	1.38%	6.49%
Hispanic	S	0.91%	0.57%	0.32%	1.77%	0.97%	1.29%	5.83%
Hispanic	W	0.81%	0.34%	0.28%	1.52%	0.71%	0.99%	4.64%
Pacific Islander	NE	0.00%	3.21%	0.00%	18.59%	0.00%	0.02%	21.82%
Pacific Islander	MW	0.00%	0.00%	0.00%	9.71%	0.00%	0.03%	9.74%
Pacific Islander	S	0.00%	0.00%	0.00%	14.23%	1.90%	1.01%	17.14%
Pacific Islander	W	0.87%	0.44%	0.01%	3.02%	0.43%	0.88%	5.65%
White	NE	0.92%	0.21%	0.30%	1.37%	1.03%	0.25%	4.07%
White	MW	0.46%	0.11%	0.19%	1.30%	0.80%	0.37%	3.23%
White	S	0.62%	0.19%	0.22%	1.69%	1.25%	0.54%	4.51%
White	W	0.48%	0.12%	0.22%	2.18%	1.05%	0.57%	4.63%

Table A-25: Standard Errors of Types of Erroneous Enumerations by Domain and Region

Domain	Region	Duplicate	Fictitious	Geocoding Error	Insufficient Information	Other Residence	Unresolved	Total EE
American Indian on reservation	NE	0.00%	1.85%	0.00%	2.78%	0.00%	0.08%	4.18%
American Indian on reservation	MW	0.25%	0.04%	0.00%	0.23%	0.24%	0.44%	0.60%
American Indian on reservation	S	1.61%	0.12%	0.00%	0.13%	0.74%	0.45%	1.84%
American Indian on reservation	W	0.15%	0.04%	0.12%	0.23%	0.32%	0.13%	0.40%
American Indian off reservation	NE	0.99%	0.71%	0.00%	1.90%	0.35%	0.21%	2.29%
American Indian off reservation	MW	0.37%	0.00%	0.21%	0.49%	0.36%	0.72%	0.94%
American Indian off reservation	S	0.31%	0.17%	0.16%	0.74%	0.42%	0.34%	0.89%
American Indian off reservation	W	0.16%	0.16%	0.09%	0.89%	0.41%	0.31%	1.05%
Hispanic	NE	0.24%	0.14%	0.20%	0.20%	0.15%	0.17%	0.48%
Hispanic	MW	0.29%	0.16%	0.07%	0.25%	0.17%	0.33%	0.58%
Hispanic	S	0.10%	0.12%	0.14%	0.16%	0.10%	0.14%	0.32%
Hispanic	W	0.10%	0.09%	0.12%	0.11%	0.07%	0.11%	0.25%
Black	NE	0.26%	0.14%	0.02%	0.38%	0.12%	0.10%	0.49%
Black	MW	0.16%	0.32%	0.03%	0.23%	0.10%	0.29%	0.49%
Black	S	0.10%	0.10%	0.20%	0.14%	0.08%	0.07%	0.30%
Black	W	0.12%	0.13%	0.17%	0.42%	0.19%	0.34%	0.58%

Table A-25: Standard Errors of Types of Erroneous Enumerations by Domain and Region

Domain	Region	Duplicate	Fictitious	Geocoding Error	Insufficient Information	Other Residence	Unresolved	Total EE
Pacific Islander	NE	0.00%	3.31%	0.00%	10.02%	0.00%	0.02%	10.35%
Pacific Islander	MW	0.00%	0.00%	0.00%	4.82%	0.00%	0.03%	4.83%
Pacific Islander	S	0.00%	0.00%	0.00%	5.39%	1.93%	0.96%	5.31%
Pacific Islander	W	0.23%	0.24%	0.01%	0.85%	0.18%	0.23%	1.00%
Asian	NE	0.53%	0.23%	0.03%	0.33%	0.24%	0.26%	0.71%
Asian	MW	0.28%	0.24%	0.00%	0.40%	0.31%	0.26%	0.64%
Asian	S	0.21%	0.14%	0.79%	0.23%	0.20%	0.18%	0.87%
Asian	W	0.26%	0.10%	0.11%	0.26%	0.14%	0.10%	0.44%
White	NE	0.09%	0.04%	0.09%	0.07%	0.06%	0.02%	0.16%
White	MW	0.04%	0.02%	0.07%	0.06%	0.04%	0.03%	0.12%
White	S	0.04%	0.04%	0.07%	0.06%	0.06%	0.03%	0.13%
White	W	0.05%	0.02%	0.06%	0.10%	0.08%	0.04%	0.17%

Appendix B

Technical Documentation

The following files were used in the creation of this memo:

- Pre-Collapsed Post-Stratum Summary File (Haines, February, 2001)
- E-Sample Person Dual System Estimation Output File (Haines, February, 2001)
- HCEF (Philipp, April, 2001)
- Sample Design File (Cromar, June, 2000)
- PerMARCS Account File (Jones, December, 1999)
- HDF
- HCUF (Philipp, February, 2001)

Variable	Variable Name	File	Collapsing
Race/Hispanic Origin	domain	DSE	
Age/sex	agesex	DSE	
Tenure	tenure2	DSE	
Mail Return Rate	rrateind	DSE	
Place Size and TEA	msatea	DSE	
Region	region	DSE	
Response Method	pft, pcmode, rhhmem	HCUF	enumerator: pcmode = 2 or pft=5, 6, 17, 18 non-proxy: enumerator & rhhmem=0, 1 proxy: enumerator & rhhmem=2, 3
Imputation	amtimp	DSE	
HU match status	afnmat_hu	DSE	Corr Enum: afnmat_hu=CE, M Err Enum: afnmat_hu=EE, GE, DE Unresov: afnmat_hu=MU, UE No Match: afumat_hu is missing
Conflicting HH	addcde	DSE	
Number of Units at the Basic Street Address	UBSA	HDF	1 unit: ubsa=1 2-9 units: 1 < ubsa<10 10+ units: ubsa>9
Whole/Partial Match Code	wpfin	DSE	

Form Length	pft	HCUF	Short: pft is odd Long: pft is even
NRFU	nru	HCEF	Not in NRFU universe: nru=0, 1, 2 In NRFU universe: nru=3, 4
Coverage Edit Follow-up	ceu	HCUF	No coverage edit follow-up: ceu=0 Coverage edit follow-up: ceu=1, 2, 3, 4
Coverage Improvement Follow-up	ciu	HCUF	No coverage improvement follow-up: ciu=0, 1 Coverage improvement follow-up: ciu=2, 3, 4, 5, 6, 7, 8, 9
Capture Method	rcapsys	HCUF	Scanned: rcapsys=1 Not scanned: rcapsys=0, 2, 3
Data Capture Center	dcc	HCUF	
A.C.E. person follow-up	fuflag	DSE	No Follow-up: fuflag=0, missing Follow-up: fuflag=1, 2
A.C.E. Relisted Cluster	relist	Account	
A.C.E. Regional Office	lco	DSE	acero: first 2 digits of lco

Estimated E-Sample Components: Person records were placed into the appropriate categories based on the following E-Sample Person Dual System Estimation Output File (Haines, 2001) variables:

- ▶ The person record's final match code (FINMAT).
- ▶ The person record's initial probability of correct enumeration (CEPROBI). This does not include the person's probability of being duplicated to a surrounding block.
- ▶ The person's final probability of correct enumeration (CEPROBF). This includes the person's probability of being duplicated to a surrounding block
- ▶ The code that used A.C.E. person follow-up information to categorize unresolved people into four different categories: Fictitious Records, Geocoding Error, Other Residence and Remaining unresolved (FU_CODE2). This was used to impute the person's probability of being a correct enumeration.

- The final E-Sample person weight that reflects the probability of selection in all stages of sampling including Targeted Extended Search sampling, noninterview adjustment and weight trimming (TESFINWT).

The probability of erroneous enumerations was obtained as follows:

Type of Erroneous Enumeration	Probability of erroneous =1	Probability of erroneous between 0 and 1
Duplicate:	FINMAT=DE	FINMAT=CE, M, MR, MU, UE the probability of duplication to surrounding block is CEPROBI-CEPROBF
Fictitious Records:	FINMAT=FE	
Geocoding Error:	FINMAT=GE	
Other Residence:	FINMAT=EE, MN	
Insufficient Info	FINMAT=KE	
Unresolved Cases ²		FINMAT=P, GU, MU, UE the probability of erroneous enumeration is 1-CEPROBI

When possible, each unresolved case was classified as a type of erroneous enumeration based on the A.C.E. person follow-up information:

Type of Erroneous Enumeration	Classification of Unresolved Cases
Fictitious Records:	FU_CODE2=1 & FINMAT=P, MU, UE
Geocoding Errors:	FINMAT=GU
Other Residence:	FU_CODE2=2 & FINMAT=P, MU, UE
Remaining unresolved	FU_CODE2=3, 4, 8, 9 & FINMAT=P, MU, UE

The Remaining unresolved cases were redistributed into two of the possible categories (Fictitious Records and Other Residence) based on the proportion of unresolved cases in each of these categories. The cases were weighted with the TESFINWT. A ratio adjustment of 265,580,677/264,578,863 was applied. This was the ratio of the E-Sample Universe to the weighted total E-Sample estimate.

²For unresolved cases the probability of being erroneously enumerated due to duplication is accounted for in the "Duplicate" category.

Total Erroneous Enumerations: was obtained by summing the Duplicates, Fictitious Records, Geocoding Errors, Other Residence, and Insufficient Information.